

Medical University of South Carolina

MEDICA

MUSC Theses and Dissertations

Winter 12-12-2022

Developing and Utilizing Novel Biofluid N-Glycan Profiling Methods for the Classification of Suspicious Mammogram Findings

Calvin Blaschke
Medical University of South Carolina

Follow this and additional works at: <https://medica-musc.researchcommons.org/theses>



Part of the [Biology Commons](#)

Recommended Citation

Blaschke, Calvin, "Developing and Utilizing Novel Biofluid N-Glycan Profiling Methods for the Classification of Suspicious Mammogram Findings" (2022). *MUSC Theses and Dissertations*. 757.
<https://medica-musc.researchcommons.org/theses/757>

This Dissertation is brought to you for free and open access by MEDICA. It has been accepted for inclusion in MUSC Theses and Dissertations by an authorized administrator of MEDICA. For more information, please contact medica@musc.edu.

Developing and Utilizing Novel Biofluid N-Glycan Profiling Methods for the
Classification of Suspicious Mammogram Findings

by

Calvin Robert Keith Blaschke

A dissertation submitted to the faculty of the Medical University of South Carolina
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
the College of Graduate Studies.

Department of Cell and Molecular Pharmacology and Experimental Therapeutics

2022

Approved by:

Chairman, Advisory Committee

Richard Drake

Anand Mehta

Peggi Angel

Carl Atkinson

Shikhar Mehrotra

TABLE OF CONTENTS

ABSTRACT.....	vi
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
ACKNOWLEDGMENTS.....	xi
Chapter 1: Literature Review.....	1
1.1 Mass Spectrometry.....	2
1.1.1 General Principles.....	2
1.1.2 MALDI.....	4
1.1.3 Mass Spectrometers in Dissertation.....	7
1.1.3.1 Bruker Solarix 7T FTICR.....	7
1.1.3.2 Bruker timsTOF fleX.....	9
1.2 N-Glycosylation.....	11
1.2.1 Composition and Structure.....	12
1.2.2 Biosynthesis.....	16
1.2.3 Biological Functions.....	20
1.3 N-Glycan Profiling.....	24
1.3.1 Serum and Plasma.....	25
1.3.2 Urine.....	29
1.3.3 Prostatic Fluids.....	30
1.3.4 Immunoglobulin G.....	31
1.4 Breast Cancer.....	34
1.4.1 Incidence.....	34
1.4.2 Risk Factors.....	35
1.4.3 Molecular Subtypes.....	37
1.4.4 Screening.....	38
1.4.5 Biomarkers.....	41
1.4.6 N-Glycan Alterations.....	41
1.4.6.1 Tissue N-Glycan Profiling.....	42
1.4.6.2 Serum N-Glycan Profiling.....	43
1.5 Broad Overview.....	52
Chapter 2: Hypothesis.....	55
2.1 Summary.....	56
2.2 Significance.....	57
2.3 Innovation.....	58
2.4 Specific Aim 1.....	59
2.5 Specific Aim 2.....	60

Chapter 3: Serum N-glycan Profiling by MALDI-IMS.....	61
3.1 Introduction.....	63
3.2 Methods.....	66
3.2.1 Materials.....	66
3.2.2 Samples.....	66
3.2.3 Array Preparation.....	67
3.2.4 Sample Capture and Washing.....	67
3.2.5 N-Glycan Release and Matrix Application.....	68
3.2.6 MALDI Imaging Mass Spectrometry.....	68
3.2.7 Data Analysis.....	69
3.2.8 MS/MS.....	70
3.3 Results.....	70
3.4 Discussion.....	88
 Chapter 4: Urine and Prostatic Fluid N-glycan Profiling by MALDI-IMS.....	 93
4.1 Introduction.....	95
4.2 Methods.....	97
4.2.1 Materials.....	97
4.2.2 Clinical Samples.....	97
4.2.3 Control Urine Standards.....	99
4.2.4 EPS Fluids and Urine Preparation.....	99
4.2.5 Prostate Tissue Preparation.....	100
4.2.6 MALDI Imaging Mass Spectrometry.....	101
4.2.7 Data Processing and Analysis.....	101
4.3 Results.....	104
4.4 Discussion.....	118
 Chapter 5: Classifying Suspicious Mammogram Findings by Serum and IgG N-glycan Profiling.....	 122
5.1 Introduction.....	124
5.2 Methods.....	126
5.2.1 Serum Samples.....	126
5.2.2 Materials.....	128
5.2.3 Serum Analysis Sample Preparation.....	129
5.2.4 IgG Analysis Sample Preparation.....	130
5.2.5 Mass Spectrometry Data Collection.....	131
5.2.6 Data Processing.....	132
5.2.7 Statistics.....	133
5.3 Results.....	134
5.4 Discussion.....	144
 Chapter 6: Conclusions, Limitations, and Future Studies.....	 150
6.1 Overall Findings.....	151
6.2 Serum and Plasma N-glycan Profiling by MALDI-IMS.....	151

6.2.1 Conclusions.....	151
6.2.2 Limitations and Future Studies.....	152
6.3 Urine and Prostatic Fluid N-glycan Profiling by MALDI-IMS.....	154
6.3.1 Conclusions.....	154
6.3.2 Limitations and Future Studies.....	155
6.4 Classifying Suspicious Mammogram Findings by Serum and IgG N-glycan Profiling.....	156
6.4.1 Conclusions.....	156
6.4.2 Limitations and Future Studies.....	157
6.5 Final Thoughts.....	159
 REFERENCES.....	 160

ABSTRACT

CALVIN ROBERT KEITH BLASCHKE. Developing and Utilizing Novel Biofluid N-glycan Profiling Methods for the Classification of Suspicious Mammogram Findings. (Under the direction of RICHARD DRAKE).

Biofluids are a great source of biomarkers because they reflect the immune and metabolic status of cells throughout the body and can be collected non-invasively. Changes in the levels and compositions of N-glycans released from serum and plasma glycoproteins have been assessed in many diseases across many large clinical sample cohorts. Assays used for N-glycan profiling in these fluids currently require multiple lengthy processing steps, thus diminishing their potential for use as standard clinical diagnostic assays. In response to this need for rapid, simple, and high-throughput biofluid analysis, a novel slide-based biofluid profiling platform was developed for the detection of N-glycan alterations that can function as biomarkers of disease. This platform was initially validated for the analysis of serum and plasma N-glycan profiles but was also adapted for the evaluation of urine and prostatic fluid N-glycan profiles. Key to these workflows was the immobilization of the fluid glycoproteins to a slide that was washed of all contaminants, followed by the rapid and highly sensitive detection of enzymatically released N-glycans by mass spectrometry. The enzyme used to release the N-glycans can be changed to target and increase the sensitivity for specific N-glycan

classes. To demonstrate the utility and feasibility of applying this platform, serum samples from 199 women with breast cancer and 99 women with a benign lesion in their breast were analyzed in order to identify differences in the N-glycan profiles. The overall N-glycan profiles of the two patient groups had no differences, but there were several individual N-glycans with significant differences in intensities between patients with benign lesions and ductal carcinoma in situ (DCIS). For women aged 50 - 74 with a body mass index of 18.5 - 24.9, a model including the intensities of two N-glycans, 1850.666 m/z and 2163.743 m/z, age, and BMI were able to clearly distinguish the breast cancer patients from the patients with benign lesions with an AUROC of 0.899 and an optimal cutoff with 82% sensitivity and 84% specificity. This study indicates that serum N-glycan profiling is a promising approach for providing clarity for breast cancer screening, especially within the subset of healthy weight women in the age group recommended for mammograms. Overall, the workflows described in this dissertation have displayed the sensitivity, adaptability, and throughput to be utilized as a biomarker discovery platform with significant clinical utility.

LIST OF TABLES

Table 1. N-glycans observed in serum and plasma.....	74
Table 2. N-glycans observed in urine and prostatic fluids.....	103
Table 3. Clinical characteristics of patients by diagnosis.....	128
Table 4. N-glycans observed in serum samples of patients.....	135
Table 5. Serum N-glycans with significant differences between benign and DCIS patients.....	137
Table 6. Serum N-glycans with significant associations with age.....	139
Table 7. Serum N-glycans with significant associations with BMI.....	140
Table 8. Serum N-glycans with significant differences between benign and cancer patients of age 50-74 and BMI 18.5-24.9.....	141
Table 9. N-glycans observed in IgG from serum samples of patients.....	143

LIST OF FIGURES

Figure 1. A schematic diagram of the MALDI process.....	5
Figure 2. A schematic diagram of a typical ICR cell geometry.....	9
Figure 3. The symbols, chemical structure, and weight of the common monosaccharides found in human N-glycans.....	13
Figure 4. Representations of the N-glycan classes.....	15
Figure 5. Breast cancer clinical characteristics by stage at diagnosis.....	40
Figure 6. Workflow for slide-based serum and plasma N-glycan profiling by MALDI IMS.....	73
Figure 7. Representative spectra of a standard serum sample and a standard plasma sample analyzed with the N-glycan profiling method by MALDI FTICR MS.....	76
Figure 8. Repeatability of the N-glycan profiling method for serum analysis.....	77
Figure 9. Repeatability of the N-glycan profiling method for plasma analysis.....	79
Figure 10. Overview of the enzymatic treatments applied for serum analysis.....	80
Figure 11. Endo F3 modification of the N-glycan profiling method for serum analysis.....	81
Figure 12. MS/MS of serum N-glycans by CID done directly on the slide.....	82
Figure 13. Sialidase modification of the N-glycan profiling method for serum analysis.....	84
Figure 14. Alternative enzymatic treatments increase serum analysis sensitivity to classes of N-glycans.....	85
Figure 15. Application of the N-glycan profiling method to the serum of breast cancer patients.....	87
Figure 16. Workflow for slide-based urine and prostatic fluid N-glycan profiling by MALDI IMS.....	105
Figure 17. SDS gel separation of the samples.....	106
Figure 18. N-glycan profile of prostatic fluids analyzed by MALDI IMS.....	108
Figure 19. Mass spectra of representative EPSu and EPSd samples with annotations for a set of putative N-glycan structures.....	112
Figure 20. N-glycan profile of pooled healthy male and female urine analyzed by MALDI IMS.....	113
Figure 21. N-glycan imaging of FFPE prostate cancer and non-cancer tissues.....	114
Figure 22. Higher resolution images of the H&E staining of each slide and scanning into a Hamamatsu.....	115
Figure 23. Detection of intraluminal N-glycans in non-tumor prostate glands.....	116
Figure 24. Comparison of detected N-glycans across sample sets.....	117
Figure 25. Workflow for slide-based IgG N-glycan profiling by MALDI IMS.....	131
Figure 26. Sample collection, processing, and analysis.....	134
Figure 27. N-glycan class intensities of serum samples with benign lesions, DCIS, and invasive breast cancer.....	136
Figure 28. Serum N-glycan differences between benign, DCIS, and invasive samples.....	138

Figure 29. Correlation of serum N-glycan intensities with the patient age and BMI.....	140
Figure 30. Serum N-glycan differences between benign and cancer samples from patients aged 50 – 74 and with a BMI of 18.5 – 24.9.....	142
Figure 31. N-glycan class intensities of IgG from serum samples with benign lesions, DCIS, and invasive breast cancer.....	144

ACKNOWLEDGMENTS

There are many people that I need to thank because without them I would not have been able to complete the work described in this dissertation. First, I would like to thank Dr. Drake for his mentorship. I have always felt that I had your support throughout completing my degree. I have appreciated your input on tackling the many problems that came up along the way, feedback on my development, and the countless great ideas and new directions you came up with to elevate my work. The culture you have fostered within our lab made it easy and enjoyable to come to work every day. I would also like to thank all the members of my committee. Dr. Mehta, rotating in your lab and taking the class you offered helped me form the fundamental knowledge to tackle my research question. I have also appreciated how willing you have been to share resources and materials to expediate my research. Dr. Angel, you have always been willing and eager to train me on how to use, optimize, and fix issues with the mass spectrometers. You not only taught me technical skills but have consistently provided great feedback on my writing and research approach. Dr. Atkinson and Dr. Mehrotra, thank you for all the time you have devoted to improving my training as a scientist. You both have helped me think about my research within the broader context of biologically and clinically. I would also like to thank Dr. Hill. I consider you an unofficial member of my committee, and I can not thank you enough for all the time you spent training me how to use R, explaining the most appropriate statistical approach, and tackling complex issues.

There are many other students at MUSC that have also been instrumental for my progress. Colin, you have not only been supportive of my research by being someone to run ideas by and troubleshooting, but also by providing me the motivation to be the best student possible. You were also accommodating enough to share hotel rooms with me, and I don't know many other people that can do that. Andrew, I appreciate all of the daily conversations about life, research, and lab woes. You have been a person I could always go to to work out a problem or to validate I was taking the best approach to my research. Sharon, I appreciate your continued interest in how my research was going and your help generally keeping me in the loop of things I should know. I am very grateful you ended up in Dr. Mehta's lab. Jake, it has been good to have someone else working on method development and pushing the capabilities of the lab further. Being able to talk through little technical details and give each other ideas for our research is something I have been very grateful. Denys, I have enjoyed having someone else working on breast cancer that can help keep me informed about latest findings and give insights into my results. Your positive attitude has been dearly needed some days. Grace, you have given me so much help with scheduling and fixing instrument issues that I know I would have taken a lot longer to collect my data without you. Kim, your help with learning how to use and find tools around the lab was very important for me when I joined the lab. You have continued to make the lab an enjoyable place to be for me through our chats in the morning. Connor and Alyson, during your time in the labs you

gave me valuable instruction and advice for my research and scientific career. I appreciated your guidance as senior students.

I would never have pursued a PhD or research in general if it were not for my undergraduate mentor Dr. Flaherty. Ever since I toured Coker, you invested your time and energy into helping me gain experience and figure out what a good path forward for me would be. You taught me critical basic lab techniques and established a foundation for me to do the work described here. I appreciate how you let me continually take on more responsibilities in your lab so that I could grow, not only in my technical skills but in my ability to manage my time and projects while teaching other students. Gaining that kind of experience is not common for an undergraduate and helped me get into the PhD program. Your instruction in my biology classes not only gave me a firm understanding of molecular biology but inspired me to pursue research and taught me how to make complex ideas understandable. I can't thank you enough for all you have done for me, and I am certain I would not be where I am today if it were not for you.

Any progress I have made academically and as a researcher must also be attributed to people who have had a significant impact on my personal life. These are the people who have helped give me perspective and motivated me when things became difficult. I need to thank my parents for providing me the opportunity to pursue an education, encouraging me to invest into what I found interesting, and raising me to work diligently. I have appreciated all of the support and encouragement of my siblings and their respective spouses, whether it be

direct and sincere or hidden in sarcasm. I would also like to thank my Seacoast family (Katelyn, Josh, Brandi, Josh, Meagan, Justin, Caroline, Andrew, Meagan, and Kevin) for their genuine care and prayers of support.

Jenna, there are so many ways that you have been crucial to my development not only as a scientist, but as a human being. You have listened to me express my frustration at the many, many failed experiments when it was needed, but also gave me encouragement and perspective when I needed it. I am not able to fully express my appreciation for you here, but fortunately I don't have to accomplish that here. Finally, I would like to thank God for all the many blessings he has put in my life, such as being able to pursue this degree and interact with so many great people along the way. It has been a joy unraveling the mechanisms of his creation.

Chapter 1: Literature Review

1.1 Mass Spectrometry

Mass spectrometry (MS) is an analytical chemistry technique that analyzes the molecular composition of a sample by measuring the abundance of molecules identified by their physical and chemical properties. While the typical measured characteristic of the molecules is mass, by combining MS with chromatography, molecules can be separated and analyzed by an array of other properties. A diverse range of scientific fields and industries have adopted the use of MS, attesting to the utility and variety of molecular targets that can be analyzed with MS. Rapid development and improvement in instrumentation, as well as increasing integration of other analytical chemistry instruments with MS in the form of hyphenated techniques, have allowed for this wide applicability.

1.1.1 General Principles

To characterize analytes by MS, the mass-to-charge ratio (m/z) of the analyte in the sample are detected. While there is a wide range of techniques and instrumentations that can be used and combined, there are common events that occur in every MS analysis. Mass spectrometers first convert analytes from a sample into gas-phase charged molecules. This process of ionization can be accomplished with a range of techniques that vary in how much energy they transfer to the analytes. Soft ionization refers to techniques with which a low amount of energy is transferred to the analytes and the covalent bonds in the analytes are not disrupted. Hard ionization transfers a high amount of energy that

can be used to fragment analytes. For the identification of biological analytes, electrospray ionization (ESI), matrix-assisted laser desorption/ionization (MALDI), and desorption electrospray ionization (DESI) are the most used ionization techniques.¹

Once the ions have been generated, they are separated by their m/z using either static or dynamic electric fields, magnetic fields, or a combination of both in the mass analyzer. There are a variety of mass analyzers that rely on specific physical properties and will determine the mass resolution, accuracy, and range of the analysis. The needs of a study greatly impact which type of mass analyzer would be most appropriate. Often higher mass resolution and mass accuracy can be achieved, but at the expense of longer data acquisition periods and slower overall analysis times. Compatible ion sources and cost can also limit the selection of the mass analyzer needed for a study.

In the final phase of MS analysis, the analyte ions are detected and converted into a signal. This can occur through the generation of a current or a secondary ion that can be amplified. As with the other components, the type of mass detector can vary depending on what mass analyzer is used and the design of the study. Ultimately, the abundance of the analyte ions that have been separated by m/z will be measured and converted into a mass spectrum.

As this general overview has displayed, nearly every stage of MS analysis can be adapted to create the most effective approach that will provide the necessary specificity, sensitivity, and speed to investigate the analyte(s) of interest. The widespread adoption of MS can be at least partially attributed to the

variety of sample types, molecular classes, and physical properties that can be targeted. Additional discussion of the specific techniques and instruments used in this dissertation are included below.

1.1.2 MALDI

Soft ionization techniques are useful when performing exploratory analysis of the molecular composition of samples, as they will not fragment the analytes and the m/z ratio can be used to identify the analytes by accurate mass. MALDI is a soft ionization technique that has been used to analyze a range of molecular targets. To analyze a sample by MALDI, a matrix, usually a weak organic acid that absorbs UV, is cocrystallized with the sample.² Typically, the matrix is applied in excess concentrations compare to the sample. The matrix can be applied by mixing with the sample in solution, spotted on top of a dried sample, or sprayed in an even coating across a sample mounted to a slide or other surface. When the sample is irradiated with laser pulse, the matrix absorbs the light energy and converts it into heat, and both the matrix and the sample are vaporized and released from the surface.³ This process is shown in Figure 1. Additionally, the matrix acts as a proton donor and receptor once struck by the laser pulse, which will ionize a portion of the sample molecules into either negative or positive ions, respectively.⁴ MALDI minimizes the fragmentation of analytes because the majority of the energy from the laser pulse is absorbed by the matrix.

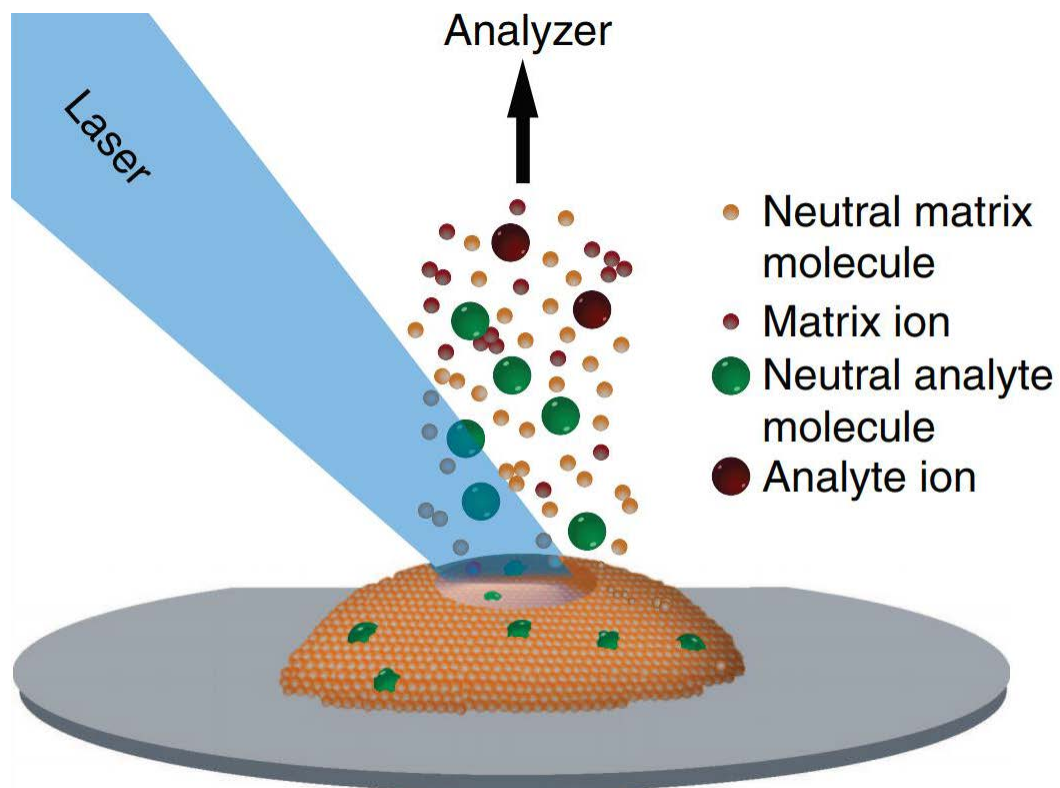


Figure 1. A schematic diagram of the MALDI process. This was adapted from Dreisewerd, K. Recent methodological advances in MALDI mass spectrometry. *Anal. Bioanal. Chem.* 406, 2261–2278 (2014).³

ESI and MALDI are some of the most used soft ionization techniques. MALDI is more compatible with complex sample mediums than ESI because it can tolerate more contaminants, such as salts and detergents.⁵ As MS measures the m/z ratio of analytes, ionization processes that create multiply charged ions, such as ESI, make analysis more complex. The ions produced using MALDI are primarily singly charged, making the m/z ratio equal to the mass of the analytes in Daltons (Da). However, MALDI often generates ions with M+H isomers and with metal adducts. These create multiple m/z ratios for a single analyte and can complicate analysis. Additionally, low molecular weight analytes can be difficult to analyze using MALDI because the matrix ions are highly abundant in the lower m/z ratio range.⁶

The ionization efficiency of the analyte is heavily dependent on the type of analyte, e.g., metabolites, drugs, lipids, peptides, proteins, or glycans, and the choice of matrix. Sinapinic acid (SA), 2,5-dihydroxybenzoic acid (DHB), and α -cyano-4-hydroxycinnamic acid (CHCA) are widely used matrices.³ Reactive MALDI matrices are being investigated for their ability to participate in a chemical reaction with the analyte during laser irradiation and improve ionization efficiency.⁷ By dissolving or suspending conventional matrices in a vacuum-stable liquid, researchers have created ionic liquid matrices (ILMs).⁸ These ILMs create a more homogenous mixture with the analytes, have a precisely controlled pH, and cause less fragmentation than traditional matrices.⁹

Historically, use of MALDI mass spectrometry for clinical diagnostic applications is frequently criticized for issues of reproducibility and laser

performance over time. Modern MALDI mass spectrometers use Nd:Yag lasers that have improved lifespans, throughput, and consistency compared to nitrogen lasers. It should be noted that in the last decade use of MALDI instruments has completely transformed clinical microbiology laboratories and clinical diagnostic workflows.¹⁰⁻¹² Key to those methods are the signature analyte barcodes used to identify individual bacterial genus species and other microorganisms.

MALDI has been coupled with Imaging Mass Spectrometry (IMS) to allow for a spatial analysis of a diverse set of biomolecules in a variety of sample types.¹³⁻¹⁵ MALDI IMS involves applying matrix to a sample and ionizing the analyte molecules with a laser pulse in a raster across the sample. A mass spectrum is generated for each coordinate, and the intensity of each ion can be visualized across the sample in a heat map. Initially developed for the analysis of fresh frozen and formalin-fixed paraffin-embedded (FFPE) tissue sections, MALDI IMS displayed great utility in linking analyte distributions with histopathological classifications of tissue subtypes by standard histology and immunohistochemistry stains.¹⁶⁻¹⁸ The sensitivity and robust nature of MALDI IMS has led to the development of novel applications of this technology for the analysis of cells and antibody-panels.^{19,20}

1.1.4 Mass Spectrometers in Dissertation

1.1.4.1 Bruker Solarix 7T FT-ICR

The novel technique described in Chapter 1 of this dissertation was initially developed using a Bruker Solarix 7 Tesla FT-ICR MS. This mass spectrometer has both an ESI and a MALDI source, 2000 Hz SmartBeam II ultraviolet laser, and utilizes an ion cyclotron resonance (ICR) cell. After ionization in the MALDI source, DC and RF voltages are used to focus the ions that have desorbed from the sample into an ion beam. The ion beam is guided from a low vacuum environment to the ultra-high vacuum analyzer by moving through a quadrupole, which will filter out any ions not within a user-specified m/z range, and a collision cell, which can be used to fragment specific ions for MS/MS analysis. Before entering the ICR cell, optical lenses are used to focus the ions. The ICR utilizes three sets of electrodes that are within a 7 Tesla magnet to analyze the ions. The magnetic field confines the ions to the cell and induces a cyclotron motion. The trapping plates generate a small electrical field to prevent the ions from escaping the ICR. The excitation plates use a RF frequency sweep to increase the radius of the ions' motions. The detection plates register the charge induced by the passing ions. A schematic of the ICR cell can be seen in Figure 2.

The ion's motion in a circular orbit creates a signal in the form of a sinusoidal wave, commonly referred to as a transient, with a periodicity related to the ion's m/z ratio. Longer transients, that is the time that the ions oscillate in the

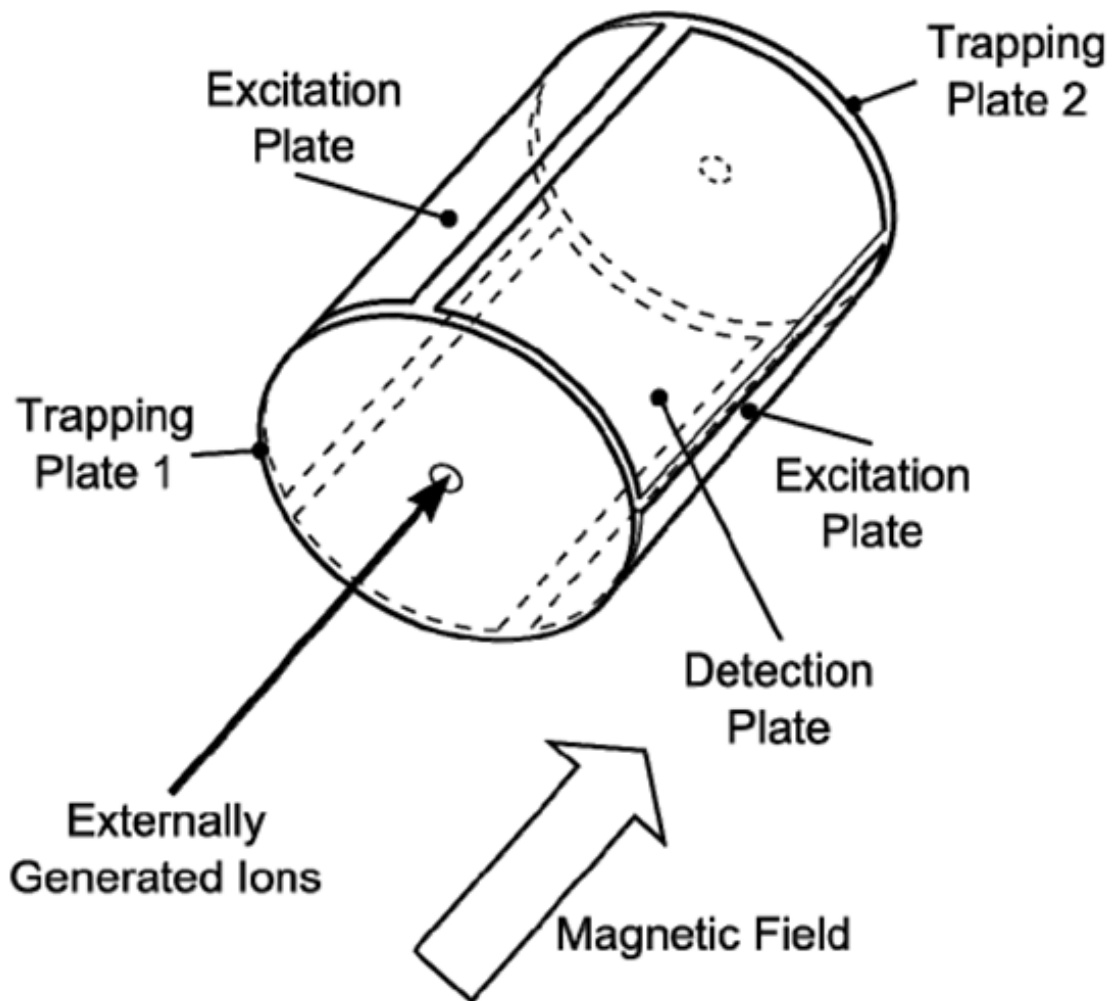


Figure 2. Schematic of a typical ICR cell geometry. This figure was adapted from Bruker Daltonics Inc. SolariX User Manual. Revision 3 (2017).²¹

ICR cell, will make the peaks of the detected ions narrower and increase the resolution of the data collected. As there are typically a mixture of ions and their naturally occurring stable isotopes with discrete m/z ratios that are being detected simultaneously, this complex time domain waveform is converted into a frequency domain spectrum, which corresponds with the m/z ratios, by a Fourier transform. This form of analysis, Fourier transform ion cyclotron resonance (FT-ICR), has high mass accuracy and the Solarix's collision cell allows for MS/MS validation of target analytes.

1.1.4.2 Bruker timsTOF fleX

The bulk of the data generated and presented in this dissertation was collected with a Bruker timsTOF fleX MS. This instrument has a dual MALDI/ESI source with a SmartBeam 3D laser operating with a 20 μm spot size and at 10,000 Hz. The timsTOF fleX also contains a collision cell that allows for MS/MS analysis. This instrument has the capability to perform ion mobility spectrometry, where the ions are separated and characterized by their movement through a buffer gas in the presence of an electric field.²² This allows for analysis of ions not only based on their m/z , but their three dimensional structure. This type of analysis is particularly helpful for differentiating isomers or analytes with small differences in molecular weight. The ion mobility functions of this instrument has not been utilized in the studies discussed in this dissertation, as there were limits in software capabilities, instrument performance, and optimizations of integrating

this type data. Instead, the timsTOF fleX was operated as a standard quadrupole time-of-flight MS.

In this mode the ions are propelled by a static electric field, which provides greater acceleration to molecules that are lower in mass or higher in charge, into a field-free region of a drift tube. The time that it takes for the ion to reach the detector at the end of the drift tube is directly proportional to the velocity of the ion, which itself is dependent on the m/z of the ion. The process of measuring the ions time-of-flight is shorter than measuring the cyclotron resonance. This form of time-of-flight (TOF) analysis results in a faster data acquisition speed for the timsTOF fleX MS compared to the Solarix 7T FT-ICR MS, but at the expense of lower mass resolution.

1.2 N-Glycosylation

The metabolic utilization and interconversion of saccharides (also called glycans) are an essential component of all living organisms. Glucose is used as an energy source and ribose and deoxyribose sugars are key components of the nucleic acid genetic code found in every organism ranging from the simplest prokaryote to the most complex eukaryotes. Complex biochemical arrangements of glycans as polysaccharides serve as structural components in plants, fungi, and animals. The attachment of a single saccharide or chain of saccharides to another class of molecule is often relevant to the function of that molecule and found in Eukaryota, Bacteria, and Archaea.^{23,24} This process, called glycosylation, is one of the most common post-translational modifications in

humans.²⁵ There are many different classes of glycosylation that are categorized based on the composition of the glycans, the composition of the non-glycan components, and the attachment site of the glycan. Examples of glycans are glycosphingolipids, glycosaminoglycans, glycosides, O-linked glycans, and N-glycans.

N-glycans are non-linear oligosaccharides covalently bound to proteins at the amide nitrogen of asparagine residues, and the process for synthesizing and attaching N-glycans is called N-glycosylation. Being the most common form of glycosylation, the structure, synthesis, and biological functions of N-glycans have been studied extensively.²⁵ As the studies detailed in this dissertation analyze samples from human patients, the background information provided here will focus on the glycosylation processes and structures that occur in humans unless stated otherwise, although much of this discussion will translate to other vertebrates.

1.2.1 Composition and Structure

Fully processed human N-glycans are comprised of primarily N-acetylglucosamine, mannose, galactose, N-acetylneuraminic acid, and fucose. Additional monosaccharides may be present in humans or other vertebrates, such as N-acetylgalactosamine, N-glycolylneuraminic acid, and glucose. The hexose monosaccharides are mannose, galactose, and glucose. They are isomers of a 6-carbon sugar with a molecular weight of 180.063 Da. N-acetylglucosamine and N-acetylgalactosamine are N-acetylhexosamines, 6-

carbon sugars with an acetyl group, and have a molecular weight of 221.090 Da. Fucose is a deoxyhexose, a 6-carbon sugar with one less hydroxyl group, and has a molecular weight of 164.069 Da. The sialic acids are N-acetylneuraminic acid and N-glycolylneuraminic acid, which weigh 309.106 Da and 325.101 Da respectively, and are charged monosaccharides. Since N-glycolylneuraminic acid is not found in humans and is not investigated in any of the studies included in this dissertation, any use of the term “sialic acid” is referring to N-acetylneuraminic acid.

As representing N-glycans with entire chemical structures becomes unproductive with larger structures, the glycobiology field has taken to using colored symbols to represent the monosaccharides.^{26,27} The symbol's and other relevant information of the common monosaccharides present in human N-glycans are displayed in Figure 3. The monosaccharides in N-glycans are covalently bound by glycosidic linkages. This results in the loss of water and correspondingly, decreases the weight of the monosaccharide by 18 Da when added to the N-glycan.

N-glycans can have a variety of structures, but all contain a base structure of two N-acetylglucosamine residues and three mannose residues. The broadest categorizations of N-glycans are based on what monosaccharide residues are present other than the core structure. These broad classes are high-mannose, hybrid, and complex (Figure 4 A). High-mannose structures have only mannose residues. Hybrid structures have at least one mannose residue and one N-acetylglucosamine residue. Complex residues have at least one N-

acetylglucosamine residue and no mannose residues. Fucosylation, galactosylation, and sialylation are N-glycan classes defined by the presence of fucose, galactose, and sialic acid residues, respectively (Figure 4 B). N-glycans can have antenna, which are a series of linked monosaccharides starting with a N-acetylglucosamine attached to the core structure. N-glycans with two antenna, three antenna with one antenna on each mannose residues, three antenna on the outer most mannose residues, and four antenna are in the biantennary, bisecting, triantennary, and tetraantennary, respectively (Figure 4 C). While N-glycans with other structures are present in humans, such as polylactosamines and alpha-galactoses, the N-glycan classes displayed in Figure 4 are the most common.

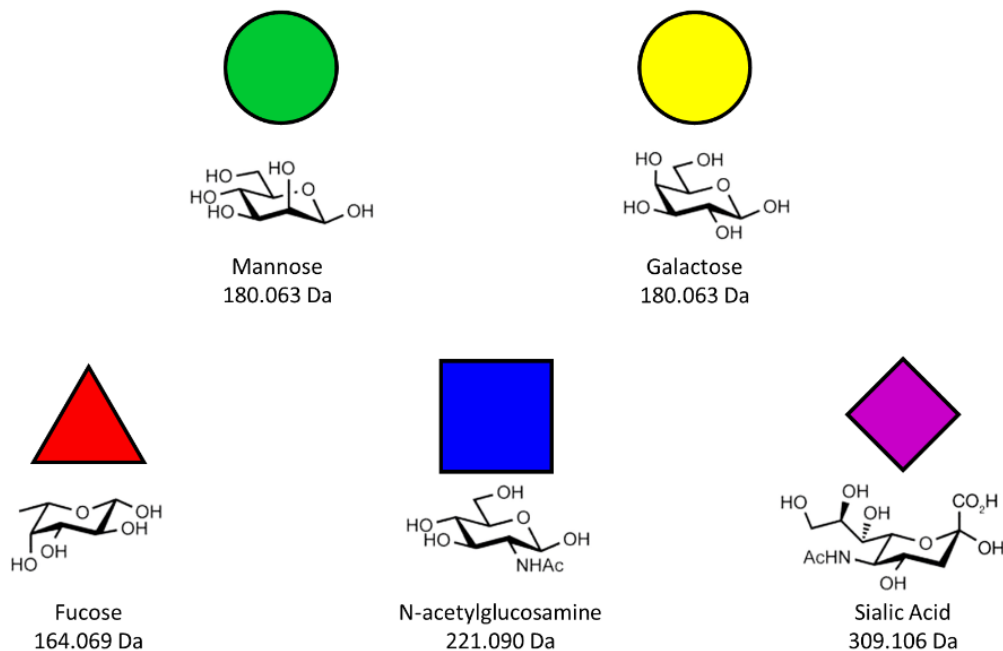


Figure 3. The symbols, chemical structure, and weight of the common monosaccharides found in human N-glycans.

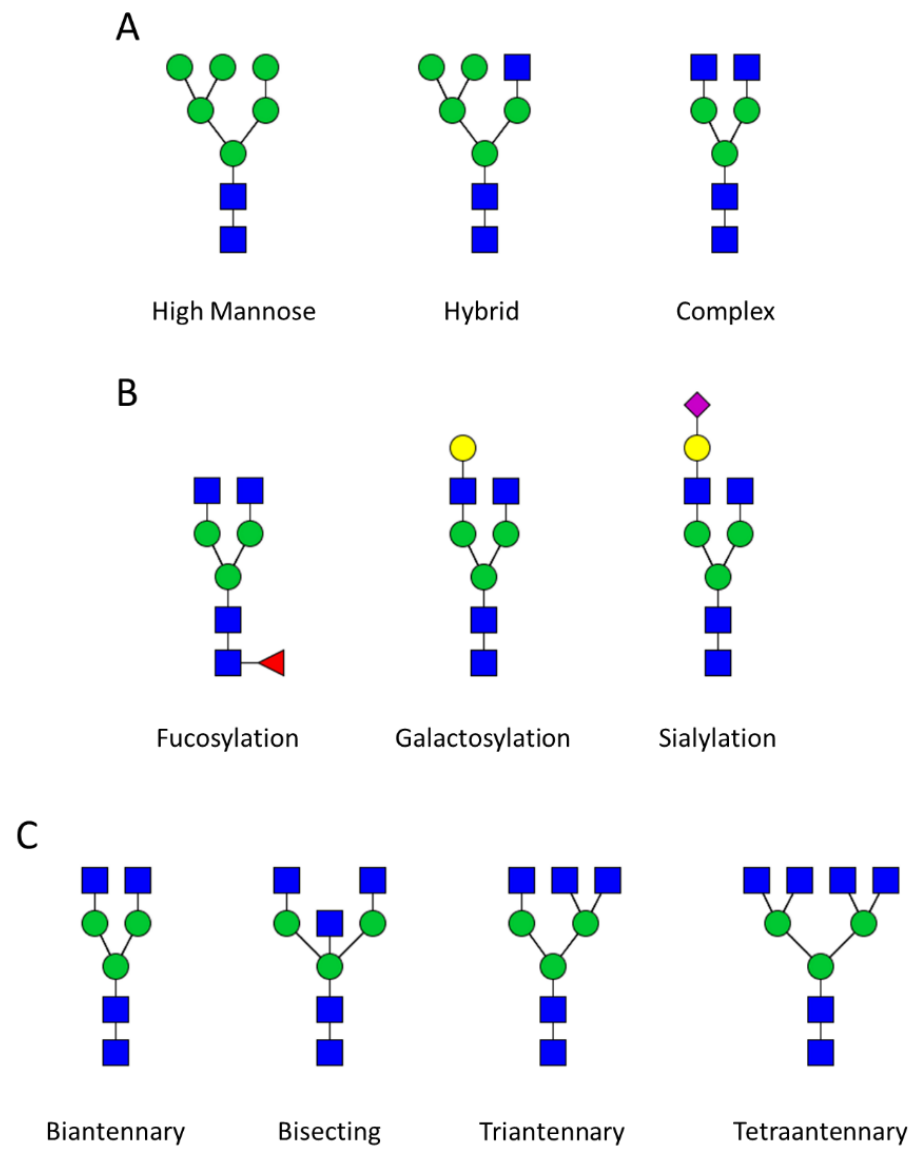


Figure 4. Representations of the N-glycan classes. A) N-glycan classes based on the presence of mannose, N-acetylglucosamine, or both on the core structure, B) the presence of specific monosaccharides, and C) the number of N-acetylglucosamines attached to the core structure.

1.2.2 Biosynthesis

N-glycan biosynthesis is a well-defined, yet complex process involving a series of enzymatic modifications to a N-glycan precursor attached to a protein in the endoplasmic reticulum (ER) and further modified in the Golgi. The enzymes that add and remove monosaccharide residues are called glycosyltransferases and glycosidases, respectively. For monosaccharides to be incorporated into an N-glycan, they must be in the form of a high-energy donor. The monosaccharides can be either salvaged from degraded glycans, imported into the cell, or enzymatically converted from other sugars within the cell. A nucleoside triphosphate and a monosaccharide with a phosphate are used to create nucleotide sugar donors which the glycosyltransferases can add via a glycosidic bond to a N-glycan structure. The nucleotide sugar donors typically used in human N-glycans are UDP-Glucose, UDP-Galactose, UDP-N-acetylglucosamine, GDP-Mannose, GDP-Fucose, and CMP-Sialic acid.

There is a complex array of N-glycan structures, and accordingly, glycosyltransferases are a large family of enzymes. The donor and acceptor substrates for glycosyltransferases are typically very specific, and glycosyltransferases will catalyze a bond between specific carbons of the monosaccharides. Glycosidases catalyze the breaking of the glycosidic bond between two specific monosaccharides either within (called endoglycosidases) or at the terminal end of the N-glycan (called exoglycosidases) and are important for the degradation of N-glycan structures for uptake and metabolism of sugars, turnover of glycoproteins, and even, regulating protein folding.

The N-glycan structure is initially constructed on a lipid carrier, a dolichol phosphate, in the membrane of the ER. An oligosaccharide with two N-acetylglucosamines, nine mannoses, and three glucoses is formed on the dolichol phosphate. The precursor is partially formed on the cytoplasmic side of the ER and then translocated across the ER membrane by a flippase where it is further processed. A protein complex called oligosaccharyltransferase (OST) transfers the N-glycan precursor to an asparagine residue on a polypeptide that is being translated and entering the ER.

Once the N-glycan has been attached to the protein, the trimming and addition of the glucose residues on the N-glycan precursor are incorporated into the cell's quality control of the protein's folding. Calnexin (CNX) and calreticulin (CRT) are ER chaperone proteins that prevent incompletely folded glycoproteins from exiting the ER by binding to the N-glycan precursors. If the glycoprotein fails to fold correctly after multiple iterations of this process, the precursor N-glycan will be truncated into a specific structure that will mark the protein for transportation to the cytoplasm where it will be degraded by a proteasome.

In the Golgi, the N-glycan will be progressively modified by glycosyltransferases and glycosidases to create a mature N-glycan with features distinctive of the N-glycan classes. The N-glycan enters the Golgi as a high-mannose N-glycan with nine mannose residues. At this point, the glycoprotein may travel through the Golgi with no further modifications to the N-glycan. As the glycoprotein is transported through the *cis*-Golgi mannose residues can be trimmed off and N-acetylglucosamine residues can be added to initiate the

formation of antenna or hybrid N-glycans. Additional N-acetylglucosamine residues and fucose residues can be incorporated in the *medial*-Golgi, and galactose and sialic acid residues can be added in the *trans*-Golgi. At this stage, N-glycans can also be modified by the addition of sulfate and phosphate groups. Sulfotransferases will catalyze the transfer of a sulfate group from 3'-phosphoadenosine-5'-phosphosulfate (PAPS) to primarily galactose or N-acetylglucosamine residues.²⁸ Phosphorylation of N-glycans occurs through a phosphate being transferred from a UDP-N-acetylglucosamine donor to a mannose residue on an N-glycan. This is primarily used to select, segregate, and traffic lysosomal enzymes through the Golgi to lysosomes.²⁹ Once the N-glycan is fully processed in the Golgi, the glycoprotein is transported to the cell membrane for incorporation or secretion.

A glycoprotein can have a set of glycoforms that are defined by the presence, structure, and composition of the N-glycans present. There are many complex and interdependent factors that influence the number and relative abundances of the glycoforms for a glycoprotein. It is important to note that while the biosynthesis of N-glycans has been described here as segmented and sequential, multiple processing steps can occur simultaneously, in a competitive manner, or in a dependent manner. Some of the factors that influence the large range of glycoforms are the concentration of sugar nucleotide donors, the localization and abundance of the glycotransferases and glycosidases, and the amount of time the glycoprotein spends in the ER and Golgi.

The variability of N-glycans that can be detected at a single site, known as N-glycosylation microheterogeneity, can be greatly impacted by the metabolic state of the cell. The hexosamine biosynthetic pathway, which synthesizes UDP-N-acetylglucosamine, is sensitive to the concentrations of UTP, glucose, glutamine, and acetyl-CoA.³⁰ This links the presence of several nucleotide sugar donors to a cell's metabolism of the four major macromolecules. As many of the monosaccharides used to build N-glycans are also used in other forms of glycosylation, the different processes are intricately linked as well.

The complexity and heterogeneity of N-glycosylation is present even at a single N-glycosylation site. N-glycans are attached to asparagine residues within an asparagine-X-serine/threonine sequon, where X is any amino acid except proline, but even when this sequon is present, it may not be N-glycosylated. It is estimated that approximately one third of the sequons are not N-glycosylated.³¹ This variability in site occupancy is called N-glycosylation macroheterogeneity, and it can be influenced by several factors. Asparagine-X-Threonine sites are more efficiently N-glycosylated than Asparagine-X-Serine sites.^{32,33} Acidic amino acids in the X position can decrease site occupancy, while an aromatic amino acid two residues before the asparagine can increase site occupancy.^{34,35} The secondary and tertiary structure of the protein also impact the efficiency of N-glycosylation. Sequons located in flexible domains of the protein are more likely to be N-glycosylated, but a majority of the sequons are found in β -sheets.^{35,36} OST can have two different catalytic subunits that either glycosylate polypeptides as they enter or after they enter the ER. If a N-glycosylation site is not

glycosylated as the polypeptide enters the ER, the tertiary structure of the protein can affect whether the N-glycosylation site is accessible for OST. Even if a N-glycan is added to a protein in the ER, the tertiary structure of the protein can inhibit glycan-processing enzymes in the Golgi from accessing the N-glycan.

Regulation of glycosyltransferases and glycosidases can influence the N-glycan profiles of a cell's glycoproteins. Transcriptomic analysis by large-scale microarray analysis identified distinct N-glycosylation processing enzymes for different cell types.^{37–39} Additionally, to be fully active many of these enzymes must be post-translationally modified as well by either phosphorylation or glycosylation.^{40,41} Specific N-glycan biosynthetic enzymes involved in creating high mannose and fucosylated structures are regulated by microRNA at the transcriptional level.⁴² The biosynthesis of N-glycans is a heavily regulated process that can be influenced by numerous cellular processes and environmental cues. This allows for a dynamic modulation of the glycoforms present on the glycoproteins and creates a diverse set of N-glycan structures.

1.2.3 Biological Functions

The impact of N-glycosylation can be identified early in the development of an organism. The different developmental stages of vertebrates have been linked to distinct N-glycan profiles.⁴³ Genetic modifications that affect the early stages of N-glycan processing are often embryonic lethal.⁴⁴ Mutations that affect enzymes involved in adding sialic acid residues often have less dramatic effects, since sialylation occurs in the later stage of N-glycan processing and affects only the

terminal ends of the structures.⁴⁵ The rarity of human congenital disorders of glycosylation (CDGs), inherited human diseases that are caused by defects in glycan biosynthesis and metabolism, highlight the importance of glycosylation in embryonic development. CDGs can have highly variable impacts, even in the same mutation, but usually affect multiple organ systems.⁴⁶

The linkages, or the carbons that form the glycosidic bond between the monosaccharides, of an N-glycan can impact its effect on its carrier protein. Some of the most well-documented linkage specific effects are centered around sialic acids and fucoses. Fucose residues can be bound to the asparagine linked N-acetylglucosamine in a α 1,6 linkage or the N-acetylglucosamine on the N-glycan branches in a α 1,3 or α 1,4 linkage. These are called core fucosylation and outer arm fucosylation, respectively. Sialic acids can be attached to a terminal galactose by a α 2,3 or α 2,6 linkage.

N-glycans have important effects on proteins at the structural level due to their hydrophilic and flexible chemical nature. As previously described, N-glycan binding with calnexin and calreticulin play integral roles in assuring correct protein folding before the protein is secreted or integrated into the plasma membrane. The presence of the core N-glycan structure has been shown to improve both the kinetics and thermodynamics of protein folding.⁴⁷ The presence of these hydrophilic, bulky, noncharged N-glycan structures aids in keeping the protein in solution during folding. N-glycans can facilitate oligomerization or stabilize oligomers through interactions with the protein subunits, and can even aid in the fibrilization of prion proteins.^{48,49} Investigation of aquaporin-2 found that

N-glycans stabilize mutated versions of the protein in the ER, allowing for more time for the protein to fold in a manner that minimizes the effect of the mutations instead of being targeted for degradation.⁵⁰ Site-directed mutagenesis of N-glycan sites of a bile acid transporter, ASBT, resulted in an increased sensitivity to proteolysis and decrease in function when the protein was unglycosylated.⁵¹

The N-glycans of a protein can also impact the localization, enzymatic activity, substrate specificity, and binding affinity.⁵² Placenta bikunin (PB) is a serine protease inhibitor with two glycoforms that determine its subcellular distribution and consequently, access to target proteases.⁵³ Receptor oligomerization, ligand binding, and signaling can be significantly altered by a change in or lack of N-glycans. Tumor necrosis factor- α receptor 1 (TNFR1) is a key mediator in the inflammatory-activation of microglia which contributes to the pathogenesis of several central nervous system (CNS) diseases. N-glycosylation of TNFR1 did not impact the intracellular localization, but has been found to be integral to the ligand-binding affinity and downstream signaling.⁵⁴ Epidermal growth factor receptor (EGFR) is an important mediator for cell proliferation signaling pathways. In order for EGFR to dimerize and bind growth factors, there have to be specific strong non-covalent interactions near the growth factor binding site between the N-glycans and the extracellular domain of EGFR.⁵⁵

N-glycans play significant roles in the development and functioning of multiple physiological systems. Murine embryonic stem cells that lack complex N-glycans were unable to form bronchial epithelium.⁵⁶ Mice incapable of fucosylating the N-acetylglucosamine closest to the reducing end of the N-glycan

could not develop fully functional lungs and displayed emphysema-like symptoms.⁵⁷ Glucose-stimulated insulin secretion depends upon N-glycans maintaining glucose transporters in the proper orientation.⁵⁸ Correspondingly, the N-glycosylation of pancreatic beta cell glucose transporter-2 (Glut-2) is altered in type 2 diabetes.⁵⁹ Thyroglobulin is the protein backbone for synthesis of thyroid hormones, and is modified with N-glycans that aid in the transport of thyroglobulin into the follicular lumen of the thyroid.⁶⁰ N-glycosylation, specifically galactosylation and sialylation, is dynamically regulated in the myogenesis and has been experimentally manipulated in mice to increase muscle mass.⁶¹ The prevalence of cognitive and/or neurological abnormalities in patients with CDGs are indicative of the role of N-glycosylation in neural development and function.⁶² By inhibiting the formation of complex or hybrid N-glycans, a study found that the sodium channels and electrical signaling were altered in cardiac tissue and that these changes accelerated the progression of heart disease.⁶³

The glycoproteins on the surface of cells in the immune system work with glycan-binding proteins to sense environmental signals. B cells and T cells heavily express the glycoproteins CD43 and CD45 on their surface. These N-glycosylation of these proteins are crucial for their ability to regulate cellular migration, T cell receptor signaling, apoptosis, and cell survival.⁶⁴ Galectins, soluble proteins that will bind to glycans that have galactose residues, are involved in many roles of the immune system. Galectin 1 will bind to the N-glycans of CD45 to induce apoptosis, but this interaction is dependent on the number of antenna and extent of sialylation of the N-glycans present.⁶⁵ Galectins

can also bind to receptors on the cell surface and oligomerize with other galectins to form a lattice that will cluster the receptors and influence cell-cell interactions. The result of this type of interaction is a regulation of the strength of immune cell signaling that is dependent on the N-glycan profile of different receptors. Similar to galectins, siglecs are sialic acid-binding proteins that are commonly expressed on immune cells. Once bound to an antigen, B cell receptor nanocluster formation is dependent on the action of CD22, a α 2,6-linked sialic acid-binding siglec.⁶⁶ Additionally, CD22 is essential for regulating self-tolerance as indicated by the high amounts of autoreactive IgG antibodies in CD-22 deficient mice.⁶⁷ Selectins are cell adhesion molecules that recognize sialylated and fucosylated N-glycans. Selectins are crucial for targeting leukocytes to a site of injury by binding to the leukocytes in circulation to the endothelium for extravasation.⁶⁸ As they are critical for the initiation of inflammation, selectins have even been targeting as therapeutic targets for inflammatory diseases.^{69,70} Immunoglobulins, specifically Immunoglobulin G (IgG), have significantly altered functions depending on the N-glycans present on their heavy chains. This will be further discussed in a later section.

1.3 N-Glycan Profiling

Historically, the structural complexities and non-template directed synthesis of oligosaccharides have made them more difficult to study than the other major classes of biomolecules.⁷¹ Glycobiology has relied on technological advancements to discover and explore the different classes of N-glycan

structures and their functions. The analysis of determining the totality of N-glycan structures present in a sample is called N-glycan profiling. N-glycan profiling techniques differ based on the physical and chemical characteristics of the sample.

N-glycan profiling of samples for clinical applications requires a technique with a high-throughput and a high sensitivity to robustly analyze large cohorts and identify potentially slight alterations in specific N-glycans or N-glycan classes. While analysis of intact glycoproteins and glycopeptides can link N-glycan alterations found in clinical datasets back to the carrier protein, the purification and/or data processing typically needed for these levels of analysis can hinder their suitability for clinical applications. Alternatively, complete N-glycan structures can be enzymatically released from glycoproteins using peptide-N-glycosidase F (PNGase F). PNGase F will cleave all N-glycans at the site of their attachment to the asparagine residue on the carrier protein regardless of N-glycan class. The released N-glycans can then be profiled using a variety of analytical techniques depending on the sample type.

1.3.1 Serum and Plasma

The composition of serum and plasma collected during blood draws reflects dynamic decellularized biofluids comprised of thousands of proteins and metabolites. A major component of the fluids are multiple high abundance glycoproteins produced by the liver and immunoglobulins secreted by B-cells. Extensive characterization of the proteomes of serum and plasma has indicated

a 10-fold order of magnitude concentration difference from the most abundant proteins, i.e. albumin and immunoglobulins, to the lowest concentration ones.^{72,73} Current proteomic mass spectrometry studies have identified over 4300 individual proteins present in serum, yet by percentage of abundance, 22 proteins represent 99% of the content.⁷²⁻⁷⁴ Except for albumin, which is not a glycoprotein and accounts for approximately 50% of the total protein content, the majority of the 22 most abundant proteins and next 30 are glycoproteins secreted by the liver and B-cells.

Interrogation of the glycan composition of these most abundant glycoproteins for potential clinical application to differentiate healthy from disease states has been ongoing for decades.⁷⁵⁻⁷⁷ The serum or plasma N-glycan profile is very stable in an individual over time which is important for establishing the rationale behind investigating N-glycan changes as diagnostics for disease.⁷⁸ Large-scale studies have identified certain biological factors, such as age, sex, and ethnicity, that have associated alterations in the serum or plasma N-glycan profile.⁷⁹⁻⁸³ Additionally, smoking and dietary habits, physical activity, and pregnancy can impact the N-glycans in these blood-based samples.⁸⁴⁻⁸⁶ Changes in the N-glycan profiles of serum or plasma that are associated with a disease are frequently being identified. Major metabolic diseases, immunological disorders, and many cancers have been linked to discrete changes in the levels of N-glycans and N-glycan classes.⁷⁷

Due to the variability between individuals and numerous effects of other biological factors, studies aiming to identify N-glycan profile changes due to the

presence of a disease or condition require large cohorts and a high-throughput method that can analyze the numerous samples within a reasonable timeline. The currently used methods for high throughput serum N-glycan analysis are hydrophilic-interaction ultra-high-performance liquid chromatography (HILIC-UHPLC-FLD), multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF), and MALDI-TOF- MS.⁸⁷⁻⁹¹

For HILIC-UHPLC-FLD, enzymatically released N-glycans are fluorescently labelled, typically with 2-aminobenzamide (2-AB), by reductive amination. The labelled N-glycans are then purified from the rest of the sample and reagents using HILIC, which relies upon the interaction of the multiple hydroxyl groups of the N-glycans with the amine, amide, or zwitterion-bonded silica stationary phase. The N-glycans can then be analyzed by UHPLC. The sample retention times can be calibrated on an external run of dextran ladder to obtain Glucose Unit (GU) values that have established assignments of certain N-glycans. While 2-AB has been the traditional choice of fluorescent derivatizing agent, novel derivatizing agents are actively being investigated, such as 2-(diethylamino) ethyl 4-((2,5-dioxopyrrolidin-1-yl) oxy) carbonyl amino) benzoate (named InstantPC) which improves this step greatly by lowered the sample processing time and increased the fluorescence intensity of the labelled N-glycans.⁹²

The sample preparation for xCGE-LIF is similar to that of HILIC-UHPLC-FLD. N-glycans are released from the serum or plasma glycoproteins by PNGase F. For this technique, the most commonly used reagent for fluorescent labelling

is 8-aminopyrene-1,3,6-trisulfonic acid (APTS). HILIC is commonly used for the solid-phase extraction of the N-glycans prior to xCGE-LIF during which the N-glycans are separated inside a polymer-filled capillary by their charge, size, and shape and detected by the emitted light. The migration times of the N-glycans can be annotated with established database values.

One of the largest plasma N-glycan studies analyzed over 2144 individuals by MALDI-MS to identify associations of plasma N-glycans with markers of inflammation and metabolic health.⁹³ The MALDI-MS workflow involved the release of N-glycans with PNGase F, sialic acid stabilization by ethyl esterification, purification by HILIC, incorporation of chemical matrix, and spotting onto a MALDI target plate, with many of the steps completed with an automated liquid handling robot.⁹⁴ This workflow would be further optimized by limiting derivatization side-reactions, decreasing batch-to-batch variances by changing the HILIC stationary phase, and using MALDI-FTICR instead of MALDI-TOF.⁹⁰ A non-automated form of this workflow was adapted for dried blood spot N-glycan analysis, which represents an even less-invasive collection strategy.⁹⁵ These strategies have been used to investigate a wide array of diseases, including multiple myeloma,⁹⁶ inflammatory bowel disease,⁹⁷ rheumatoid arthritis,⁹¹ and colorectal cancer.⁹⁸ These extensive studies on large sample cohorts have established N-glycan detection in plasma by MALDI-IMS as a clinically relevant assay for use in clinical diagnostic laboratories.

Another method for MALDI-TOF-MS analysis of serum or plasma N-glycans utilizes an enrichment strategy termed glycoblotting, in which beads conjugated

with hydrazide groups bind N-glycans released from serum samples.⁹⁹ Not only are the bead-bound N-glycans purified with a series of washes, but the sialic acids are stabilized by methyl esterification. The hydrazide groups with attached N-glycans are released from the bead by reduction with dithiothreitol. The enriched N-glycans are spotted directly onto a MALDI target plate, mixed with matrix, and analyzed by MALDI-TOF MS. The processing time of this workflow was reduced when combined with an automated sample-processing system that can carry out the majority of the steps.¹⁰⁰ Glycoblotting-based serum analysis has been used extensively to find many disease-specific N-glycan alterations.^{101–106} The clinical utility of this workflow is distinguished by the minimal hands-on-time.

1.3.2 Urine

Urine is an attractive clinical sample for disease biomarkers because of its non-invasive collection. Urine proteins and metabolites have been used to detect decreased kidney function, diabetes, heart failure, multiple myeloma, and multiple cancers.^{107–109} Specifically, nuclear matrix protein 22 is approved by the United States Food and Drug Administration for the detection and monitoring of bladder cancer.¹¹⁰ The function of multiple organs can be directly or indirectly associated with urine proteins and their post-translational modifications. Approximately 2,600 glycoproteins have been identified in urine, a portion of which are trafficked through extracellular vesicles (EVs).¹¹¹ Most of these proteins are filtered from the blood or secreted from the kidney or urinary tract.

While several glycoforms of urinary proteins have been investigated, there have been few comprehensive studies of the urine N-glycan profile.

Theoretically, the high-throughput N-glycan profiling methods for serum can be translated over to urine analysis, but this has rarely been done without first isolating a glycoprotein of interest. The few studies that analyze the entire urine N-glycan profile utilizes a method that can be adapted for high-throughput analysis.^{112,113} Centrifugation and molecular weight centrifuge filters are typically used to isolate the proteins from the other particles and cell debris. The isolated proteins can be reduced, alkylated, and deglycosylated still on the filter. The released N-glycans can be labelled with 2-aminobenzoic acid (2-AA) for enhanced ionization efficiency and methylamidated for purification before liquid chromatography-mass spectrometry analysis (LC-MS). The urine N-glycan profiles of healthy individuals have similar compositions to serum and plasma profiles, except for the presence of a small portion of sulfated N-glycans that are not typically detected in serum and plasma samples.

1.3.3 Prostatic Fluids

The human prostate is a small organ that links both the male urinary tract and reproductive tract organs as it encompasses the urethra and ejaculatory duct junction. It functions as male accessory sex gland that produces fluids containing glycoproteins, zinc, and citric acid. The components of these prostatic fluids have important roles in the activation and motility of sperm and the liquefaction process of ejaculate. As prostatic fluids can contain secreted proteins and shed

cells from diseased tissue, prostatic fluids have been increasingly investigated as source for biomarkers of prostate diseases.

Two glycoproteins, prostate-specific antigen (PSA) and prostatic acid phosphatase (PAP), have been assessed for decades as detection biomarkers of prostate cancer, but have limited ability to distinguish between prostate cancer and noncancer prostatic disease.^{114,115} Proteomic analysis of prostatic fluids found different expression levels of seven proteins, including PSA and PAP.¹¹⁶ Incorporation of the N-glycan profile of PSA have displayed promise in improving its performance as a diagnostic and prognostic biomarker.^{117,118} While there have been several studies of the serum N-glycan profile alterations between benign prostate hyperplasia and prostate cancer, castration-resistant and castration-sensitive prostate cancer, and indolent and aggressive prostate cancer, there have been very few studies investigated the N-glycan profile of prostatic fluids.^{119–122} Similar to the methods used for high-throughput serum N-glycan profiling, prostatic fluids can be analyzed by normal-phase high-performance liquid chromatography (HPLC) of 2-AB labeled N-glycans, MALDI-TOF of permethylated N-glycans, and multi-capillary electrophoresis of APTS-labelled N-glycans.^{123–125} From the few studies that have implemented these methodologies, decreases in several triantennary and tetraantennary N-glycans and the fucosylation N-glycan class have been associated with the prostate cancer presence and progression.^{123–125}

1.3.4 Immunoglobulin G

Immunoglobulins are a key component of the adaptive immune system found in all vertebrates that function to provide long-term defense to antigens that the host has already been exposed to.¹²⁶ There are five different classes of immunoglobulin: Immunoglobulin G (IgG), Immunoglobulin M (IgM), Immunoglobulin A (IgA), Immunoglobulin E (IgE), and Immunoglobulin D (IgD). Each immunoglobulin has a heavy chain and a light chain that forms a structure that can be broken down into the crystallizable fragment (Fc) and antigen-binding fragment (Fab). IgG is one of the most abundant in serum with 10-15 mg/ml of IgG being secreted by B cells into circulation, accounting for about 10-20% of plasma protein.¹²⁷ There are four subclasses of IgG that differ in the amino acid sequence of heavy chain and the number of disulphide bonds present.¹²⁸ The Fab domain contains the antigen-binding site that recognizes specific molecular structures of antigens. The Fc domain plays an important role for mediating the effector functions of the adaptive immune system. Once IgG has bound to a pathogen, macrophages, neutrophils, and dendritic cells can engulf and destroy the pathogen through phagocytosis. Key to this process is the recognition of the Fc domain of the IgG by Fcγ receptors on these cells.¹²⁹ Fcγ receptors can also be found on natural killer cells to induce a cytotoxic response after binding to IgG. Additionally, IgG can activate the three complement cascades: the classical pathway, the alternative pathway, and the lectin pathway. Bacteria and viruses can also be blocked from binding to cell surfaces by IgG through steric

hinderance or inhibition of pathogen conformational changes, a process called neutralization.

As is often the case for glycoproteins, IgG's function is greatly affected by its N-glycan profile. Each heavy chain has a N-glycosylation site at asparagine residue 297 in the Fc domain. The N-glycans are in the interstitial space between the heavy chains and each takes part in approximately 70 noncovalent interactions with the protein surface.¹²⁷ It is estimated the Fab domain is also N-glycosylated in 15-20% of IgG molecules.¹³⁰⁻¹³² The most abundant IgG glycoform has a biantennary N-glycan with a core fucose.¹³³ This structure is most often modified by the addition of galactose residues (50% of IgG molecules) and a bisecting N-acetylglucosamine residue (32% of IgG molecules).¹²⁷ Sialic acids are less commonly attached to the N-glycan structure due to the steric hinderance of the interstitial space and the interactions of the galactose residue with a hydrophobic pocket of the protein.¹²⁷ The many interactions of the N-glycans with the protein backbone aid in maintaining the quaternary structure of the Fc domain. Accordingly, the composition of the N-glycans can greatly influence the function of the IgG by altering the binding affinity for certain Fcγ receptors.

The presence and composition of IgG N-glycans have well-established associations with inflammatory responses. Sialylation has been linked with anti-inflammatory activity, possibly through the decreased affinity for Fcγ receptors on innate immune cells, and has been exploited for the use of therapeutic antibodies.^{134,135} The core fucose of IgG N-glycans dampens the antibody-

dependent cellular cytotoxicity (ADCC) due to interactions with the N-glycans on the Fcγ receptor III.^{136,137} A core fucose and a bisecting N-acetylglucosamine are typically not present on the same IgG molecule so their opposing roles are difficult to disentangle.¹³³ IgG N-glycans without galactoses on the non-reducing end of the antenna are generally associated with pro-inflammatory responses by activation of the complement system.^{138–141} Galactosylation of IgG has been shown to have both pro- and anti-inflammatory effects depending on the disease or condition and immune cell interactions.¹³³ Aging has been strongly linked to an increase in non-galactosylated and bisected N-glycans and a decrease in digalactosylated and sialylated N-glycans.¹⁴²

The identification of IgG N-glycan alterations in diseases have prompted the development of high-throughput techniques for large-scale biomarker discovery studies. The same methods that have been discussed for serum N-glycan profiling can be applied for IgG N-glycan profiling after IgG has been isolated, often by using protein G for affinity chromatography. IgG glycopeptides can also be analyzed using a trypsin digestion instead of using PNGase F and followed by a solid-phase extraction and ESI-MS analysis.^{143,144}

1.4 Breast Cancer

Breasts are organs that are composed primarily of adipose, connective, and glandular tissue. The glandular tissue is more abundant in females and contain lobes that contain smaller lobules that are connected via milk ducts.¹⁴⁵ Breast cancer can occur in many different areas of the breast. Some of the lower

incidence breast cancers are mucinous, medullary, and Paget's disease.¹⁴⁶ The most common locations for breast cancer are the lobules (lobular) and milk ducts (ductal).¹⁴⁶ Due to the significant differences in disease progression and prognosis, the other categorization for breast cancers is based on if the tumor is invading other tissues (invasive) or if it the tumor is not invading other tissues (in situ).

1.4.1 Incidence

In the United States and globally, breast cancer is the most commonly diagnosed cancer in females, accounting for 31% and 24.5% of new cases, respectively.^{147,148} Globally, it is even the most common cancer including both sexes.¹⁴⁷ In females, breast cancer is the second highest cause of cancer-related deaths in the United States and the leading cause globally.^{147,148} In 2020, there were over 685,000 deaths and 2.3 million new cases of breast cancer globally, and it is estimated that these numbers will increase to 1 million deaths and over 3 million new cases each year by 2040 due to population growth and ageing alone.¹⁴⁹ While breast cancer occurs in both males and females, the lifetime risk of breast cancer for a man is 1 in 833 compared to 1 in 10 for a woman.¹⁵⁰ Due to this large disparity, this dissertation will focus on breast cancer in women throughout.

1.4.2 Risk Factors

There have been many genes that have abnormal amplifications or mutations that have been associated with breast cancer, the most well-known being breast cancer associated gene 1 and 2 (*BRCA1* and *BRCA2*). The products of these genes are tumor suppressor proteins that aid in DNA damage repair. While the females that are carriers of mutated *BRCA1* or *BRCA2* are 10-30 times more likely to have breast cancer, these individuals make up only 5-10% of the all of the breast cancer cases.^{146,151,152}

Other important genes associated with breast cancer incidence are *HER2*, epidermal growth factor receptor (*EGFR*), and *C-MYC*. *HER2* expression is dysregulated in breast cancer by gene amplification and re-arrangement. The *HER2* and *EGFR* proteins are tyrosine receptor kinases that signal for cell proliferation and survival upon ligand binding. The Myc protein encoded for by *C-MYC* is a transcription factor that regulates genes whose products have key roles in breast cancer initiation and progression.

The genetic predisposition to breast cancer of females that have inherited mutated forms of these genes also makes family history an important factor in determining risk. There is some variability in the reported incidence of breast cancer in females with a family history of breast cancer. Reports can range from 10-25% of breast cancer cases being related to family-history.^{153,154} This variability is most likely due to the significant influences of ethnicity, country, family size, and environmental factors on the risk of breast cancer for an individual. Breast cancer incidence in first-degree or multiple relatives is

associated with higher risk, but to account for the multitude of factors influencing family history, models that generate a family history score have been developed.¹⁵⁴

Age is an important risk factor of breast cancer for women as well. Both in situ and invasive breast cancer incidence rates increase dramatically between women aged 20-49 and women aged 50 and older.¹⁵⁵ There are reproductive factors, age at menarche, first pregnancy, and menopause, that can be informative for an individual's risk of breast cancer, as well.¹⁵⁶⁻¹⁵⁸ Additionally, exogenous hormones delivered through contraceptives have been linked to a small increase in risk of breast cancer.¹⁵⁹⁻¹⁶¹ A higher breast density, a measure of the amount of epithelial tissue and stroma compared to fat tissue, is associated with an increased risk of invasive breast cancer in women.¹⁶² This constitutes a clinical challenge, because the fibroglandular tissue that is more prevalent in dense breasts can obscure or mimic breast cancers in mammograms, the primary breast cancer screening tool.

Other risk factors of breast cancer are related to lifestyle. Estrogen receptor pathways can be triggered by an alcohol-induced increase in estrogen-related hormones into the blood. Obesity and diets with high intake of refined sugar and both saturated and trans-saturated fats are associated with inflammatory pathways and dysregulated metabolism which increases the risk of breast cancer.¹⁶³ The strength of the associations of these risk factors can differ greatly between tumor molecular subtypes.¹⁶⁴

1.4.3 Molecular Subtypes

Breast cancer can also be subtyped by the molecular characteristics of the tumor. Some of the primary markers are the presence of hormone receptors (HR), either oestrogen receptor (ER) or progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). Triple-negative breast cancer (TNBC) are tumors that do not have ER, PR, or HER2 expression. The levels of Ki67, a nuclear protein present in all active phases of the cell cycle, is a marker of cell proliferation in breast cancer, but has little utility other than prognosis assessment in stage I or II breast cancer.¹⁶⁵ The combination of these molecules, histological features, and a 50-gene expression signature (PAM50) can be used to create subtypes that inform clinical decisions due to their differences in prognosis and treatment sensitivities.^{166,167}

1.4.4 Screening

Mammography is low-dose X-ray imaging of the breast tissue. The American College of Radiology created a system to sort the results of mammograms that is called the Breast Imaging-Reporting and Data System (BI-RADS).¹⁶⁸ BI-RADS incorporates breast composition, architectural distortion, and the shape, margin, and density of any masses or calcifications. The assessment categories range from 1-6 indicating the findings are 1-negative, 2-benign, 3-probably benign, 4-suspicious, 5-highly suggestive of cancer, and 6-known biopsy-proven cancer. There is an additional 0 category indicating that additional imaging evaluations are needed. The BI-RADS 4 category is broken down into

subcategories based on the level of suspicion for cancer, but the entire category can indicate a finding that is between >2% and <95% likelihood of cancer. The recommended follow up for a BI-RADS 4 finding is a tissue biopsy.

Increased mammography screening, the current standard detection technique, has prevented many BC deaths, but it is still limited by current rates of false positives, undetected cancers, overdiagnosis, and overtreatment.¹⁶⁹ Overdiagnosis refers to a diagnosis of cancer that would otherwise never have caused symptoms or death in a person's lifetime and overtreatment is the unnecessary use of tests and treatments.¹⁷⁰ Of all women screened by mammography, 2% are recommended to have a follow-up image or biopsy; 75% of those women do not have cancer.¹⁷¹ Additionally, 12% of screened women are recommended for follow up imaging or biopsy; 95% of these women do not have cancer.¹⁷¹ When considering how many women have a mammogram, this results in 585,000 women having unnecessary biopsies every year in the United States.¹⁷² Younger women generally have more dense breasts making diagnosing BC by mammography even more challenging.¹⁷³ The combination of false-positives and overdiagnosis can result in unnecessary biopsies, anxiety, and costs.¹⁷⁴

Magnetic Resonance Imaging (MRI) is recommended alongside mammograms in women with a high-lifetime risk, but it is underutilized, possibly due to the high cost and low specificity.¹⁷⁵ Abnormal mammogram findings can be followed-up with an ultrasound, but the modest increase in sensitivity is accompanied with a significant increase in false positives.¹⁷⁶ Given the

importance of early detection and widespread application of mammographic screening, there is a need for a more sensitive assay that could be used to supplement mammography results in order to better inform clinical decision making and limit false-positives/overdiagnosis.

The stage of breast cancer at diagnosis in the United States is displayed in Figure 5 A. The stages are determined by the areas that the cancer has spread to. In situ means that the cancer is noninvasive. Localized indicates the cancer has invaded into the breast tissue and fat. If the cancer has reached the pectoral muscle or regional lymph nodes, it was labelled regional. Distant cancers have reaches distant lymph nodes and metastasized (primarily to the liver, bones, and lungs). Figure 5 B shows the 5-year survival rates for patients based on these stages at diagnosis. The vast difference between the distant cancers to all the other stages highlights the importance of early detection and population wide screening programs.

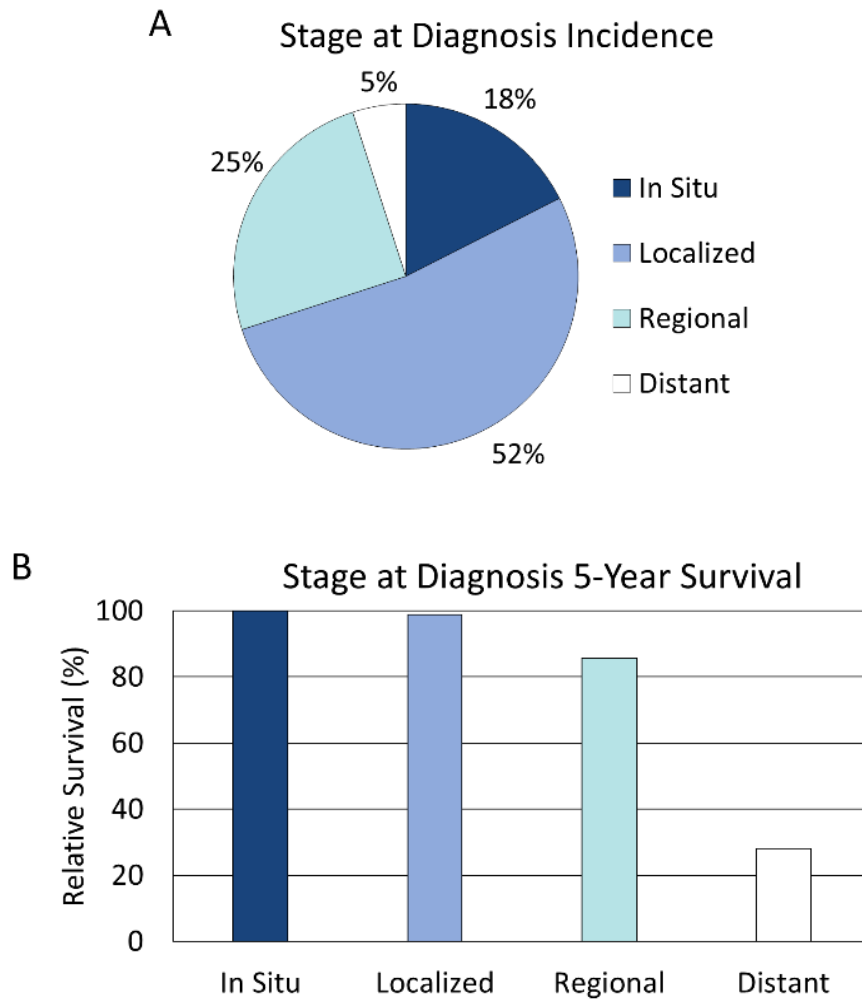


Figure 5. Breast cancer clinical characteristics by stage at diagnosis. The (A) incidence and (B) 5-year survival rates of breast cancer based on the stage at diagnosis. Data was collected from the Surveillance, Epidemiology, and End Results (SEER) program.¹⁷⁷

1.4.5 Biomarkers

Cancer antigen (CA) 15-3 has been thoroughly investigated and found to be useful for post-operative surveillance and prognosis.^{178–180} CA 15-3 detects soluble forms of a MUC1, a large heavily glycosylated transmembrane protein that has altered cellular localization and glycosylation during BC development.^{181,182} CA 27.29 is an analogous assay that binds MUC1 and has similar detection levels as CA15-3.¹⁸³ Carcinoembryonic antigen (CEA) is another high-molecular weight glycoprotein and one of the most used and recommended biomarker for BC prognosis and recurrence surveillance.¹⁸⁴ The extracellular domain of human epidermal growth factor receptor 2 (HER2) is proteolytically cleaved during BC progression and particularly effective at determining the effectiveness of trastuzumab therapy.^{185–187} Although these traditional biomarkers have utility in these roles, a lack of sensitivity and specificity when distinguishing patients with or without BC preclude CA 15-3,^{188–192} CA 27.29,¹⁸³ CEA,^{191,192} and soluble HER2¹⁹³ from effective early detection of BC.

1.4.6 N-glycan Alterations

Like many diseases, there are alterations to the N-glycosylation pathway in breast cancer. The expressions of glycosyltransferases in breast cancer tissues have been associated with certain molecular subtypes, patient prognosis, and tumor invasion and metastasis.^{194–197} Reducing expression of a sialyltransferase in a breast cancer cell line can alter the cell-cell and cell-

extracellular matrix interactions and reduce the metastatic capacity of the cancer cells.¹⁹⁸ These findings have been supported by N-glycan profiling of clinical samples.

1.4.6.1 Tissue N-glycan Profiling

Lectins are glycan-binding proteins that have a degree of specificity, typically for the terminal residues, and these proteins can be fluorescently labelled for glycan analysis. Using *Maackia amurensis* lectin (MAL), lymph node metastasis tumor samples were found to have higher expression of α 2,3-sialic acids than the matched primary tumor.¹⁹⁹ Using IMS, another group found that bone metastasis tumor samples had increased core-fucosylation and decreased high-mannose N-glycans.²⁰⁰ N-glycan profiling of benign breast tumors and breast cancer found a decrease in bisected and sialylated N-glycans in breast cancer.²⁰¹ Comparing invasive ductal carcinoma to adjacent normal breast tissue in the same patient, Yaman and colleagues identified an increase in high mannose N-glycans in the invasive ductal carcinoma tissues.²⁰² Analysis of HER2 positive and TNBC breast cancers by IMS detected differences in the abundances of a series of N-glycans in the tumor region when compared to stromal tissue.²⁰³ Additionally, this study detected poly lactosamine N-glycans in the breast cancer tissues.²⁰³

1.4.6.2 Serum N-glycan Profiling

Early studies, while not directly analyzing serum N-glycan profiles, displayed the potential for changes in the serum glycome to be indicative of BC. One of the first investigations into the changes of the glycome in BC patients utilized multilectin affinity chromatography (M-LAC) followed by high-performance liquid chromatography (HPLC) with tandem mass spectrometry (MS/MS) to isolate and identify glycoproteins, respectively.²⁰⁴ Using gene ontology (GO) terms, Yang and colleagues found increased abundances in proteins associated with cell growth and maintenance, lipid metabolism and transport, cell ion homeostasis, and a wide spectrum of protease inhibitors in the serum of patients with invasive BC. In another early study, Kirmiz and colleagues chemically cleaved glycans (both O- and N-glycans) on the glycoproteins found in the serum of BC patients and disease-free controls with sodium borohydride and identified the glycans with MALDI FTICR mass spectrometry.²⁰⁵ Using principal component analysis, the group was able to distinguish patients with BC and disease-free controls. This technique had limited use for N-glycan detection as the highest mass N-glycan was $m/z = 1442$. The findings that glycoproteins and O-linked glycans in serum are being altered in patients with BC were early indications that serum glycomics could be a valuable area of diagnostic biomarker research.

Multiple methodologies were created to investigate the serum N-glycome for BC biomarkers and overcome the challenges of N-glycan release, purification, detection, and identification in a complex biofluid. The first study to specifically investigate the N-glycan profile of BC patients' serum found changes in the

relative intensities of 8 N-glycans to be characteristic of BC.²⁰⁶ Additionally, principal component analysis of the profile data obtained from patients could distinguish healthy individuals from those in the early stages of BC. While this study analyzed permethylated N-glycans by MALDI TOF mass spectrometry, a study by Abd Hamdi and colleagues came out soon after and used normal phase (NP) and weak anion exchange (WAX)-HPLC with fluorescence detection to examine the N-glycan profiles of stage IV BC patient serum.²⁰⁷ After finding an increase in a sialyl Lewis x epitope (sLe^x) (an α 1,3-linked fucose attached to a N-acetylglucosamine at the nonreducing terminus) containing N-glycan in BC patients' serum, they isolated and quantified the triantennary, trisialylated, and outer arm fucosylated N-glycan with sequential exoglycosidase digestions and mass spectrometry. They found a 2-fold increase in this marker in BC patients compared to healthy controls and a positive correlation between the N-glycan abundance and disease progression. Additionally, they found a higher ratio of outer arm fucose compared to core fucose on the triantennary trisialylated fucosylated N-glycan for BC patients.

A chip-based reversed-phase liquid chromatography-mass spectrometry method for analyzing permethylated N-glycans was created by Alley and colleagues and used to analyze 15 stage IV BC patient serum samples.²⁰⁸ Of the 18 N-glycans detected, 5 were found to be significantly higher in the BC patients. The same group developed a protocol that selectively labels α 2,6-linked sialic acids by amidation.²⁰⁹ They utilized this protocol to examine the abundances of 12 sialylated N-glycans with linkage information between 10 cancer free serum

samples and 10 serum samples from stage IV BC patients.²⁰⁹ The additional dimension of information proved to be diagnostically useful for several N-glycans. For example, a biantennary bisialylated N-glycan was not significantly different if one sialic acid was α 2,3-linked while the other was α 2,6-linked, but it was significantly different if the sialic acids were either both α 2,3- or α 2,6-linked.

The same group that utilized chemical cleavage of glycans followed by MALDI FTICR mass spectrometry, optimized their method to allow for more sensitive detection of N-glycans and found an increase in high-mannose N-glycans in mouse serum during cancer progression and in BC patient serum.²¹⁰ In 2005, a series of lectin arrays were developed for rapid and multiplexed analysis of protein glycosylation.^{211–214} Fry and colleagues analyzed 18 serum samples, half coming from patients with nonmetastatic BC and the other half from patients with metastatic BC, with a lectin microarray consisting of 45 lectins with varying binding affinities for N- and O-linked glycans.²¹⁵ In metastatic samples they found elevated binding to *Aspergillus oryzae* lectin (AOL) and *Galanthus nivalis* agglutinin (GNA), indicative of increased fucosylation and high-mannose N-glycans respectively, and decreased binding to *Phaseolus vulgaris* erythroagglutinin (PHA-E), indicative of decreased bisected biantennary. While these methodologies had their limitations in sensitivity and sample size, the results made it clear that BC could be associated with discrete changes in the serum N-glycan profile and encouraged the development of higher-throughput methods for larger studies.

Analyzing larger data sets and using more sensitive techniques improved the potential of detecting serum N-glycome alterations associated with BC. Much of the work done during this time period was performed by the National Institute for Bioprocessing Research and Training (NIBRT) Glycoscience Group in Dublin. For one of their studies, they focused on the presence of sLe^x epitope containing N-glycans and core fucosylated N-glycans utilizing the strategy of exoglycosidase digestions, 2-AB labeling, and HPLC.²¹⁶ The serum cohort consisted of a group of 52 early BC patients' serum samples with lymph node metastasis status reported. They found that the combined abundances of these N-glycans were increased in lymph node positive BC, but if analyzed individually there was no difference. They also, and more importantly for investigating BC detection, profiled the serum of 134 patients with benign breast disease, but did not find any individual N-glycan or combination of these N-glycans able to distinguish the patients with benign breast disease or BC. This would indicate that there is no difference in sLe^x epitope containing N-glycans in the serum N-glycome of BC patients when compared to healthy controls.

The next study for this group analyzed 27 serum samples from women with advanced BC and 13 healthy controls.²¹⁷ Their main aim was to establish differences in bi-,tri-, and tetra-antennary sLe^x containing N-glycans, individually and combined, in patients with high and low circulating tumor cells. While they found an increase in each of these categories of N-glycans individually and combined (with the exception of the biantennary sLe^x containing N-glycans) in patients with high circulating tumor cells, they did not make any comment on the

differences between the healthy controls and the BC samples. Without access to the data, it is impossible to make any statistical claims but visually it seems that all of the categories of the N-glycans are higher in the high circulating tumor cell samples than in the controls. There may be a modest increase in the samples with low circulating tumor cells when compared to the controls. While this study is informative for BC prognosis or classification, the data analysis performed adds little understanding to serum N-glycan profiling's utility in BC detection.

The NIBRT group went on to utilize ultra-performance liquid chromatography (UPLC) instead of HPLC for the analysis of serum from 107 healthy women and 62 women with BC.²¹⁸ This increased their methods resolution while decreasing the run time. They detected 46 N-glycan peaks and through WAX-HPLC followed by endoglycosidase digestions, they were able to assign 165 N-glycan structures, including 22 sialic acid containing-N-glycans with differing linkage arrangements, to the peaks. Structural analysis of individual N-glycan peaks allows for determination of the predominant structure, but by UPLC analysis alone, the abundance of individual structures cannot be compared. While this method cannot directly compare specific N-glycan structures to each other, they can make comparisons between peaks and calculate differences in N-glycan features, such as fucosylation and sialylation. When analyzing the BC serum, they found a decrease in high-mannose N-glycans and core-fucosylation and an increase in sialylation, branching, and outer-arm fucosylation. The finding of decreased high-mannose N-glycans contradicts a previous study, but may be due to a more sensitive detection technique and larger sample size.²¹⁰ The peaks

containing the sLe^x epitope were increased in the BC patients and a number of peaks containing bisected N-glycans had significantly lower abundance in cancerous samples. Utilizing the serum N-glycan profiles detected, prediction analysis was able to correctly classify samples with a sensitivity and specificity of 82% each.

Ju and colleagues took a different approach than the NIBRT group and detected permethylated N-glycans using linear ion-trap quadrupole-electrospray ionization mass spectrometry (LTQ-ESI-MS).²¹⁹ For this study, 134 serum samples were donated by Chinese women (43 from healthy controls and 91 from BC patients). While they only detected 40 N-glycans, each peak represented a single N-glycan and was confirmed using multistage MS for high-confidence comparison and structural assignment. Partial least squares discriminant analysis (PLS-DA) was used to successfully discriminate samples as either from healthy controls or BC patients. All of the bifucosylated N-glycans detected were higher in the BC samples than the healthy controls and when combined for receiver-operator characteristic (ROC) analysis, had an area under the curve (AUC) of 0.93. This indicates an accurate diagnostic performance and surpassed the AUC for CEA (AUC = 0.79). Several of these bifucosylated N-glycans contained sLe^x epitopes. The high-mannose N-glycans were higher in BC patients but had low sensitivity and specificity. The increase in detection and throughput of the techniques used during this time facilitated more thorough studies with larger sample sizes.

Two recent studies have examined the serum N-glycome of BC patients, compared them to healthy controls, and found many detected N-glycans significantly altered utilizing novel techniques for the BC detection field. The first study had only Ethiopian women as subjects and analyzed 115 BC patients ranging from stage I to IV and 33 healthy controls.¹⁰² For this method, released N-glycans are captured on beads via reversible hydrazine bonds, washed, sialic-acid stabilized by methyl esterification, released with a mild acid hydrolysis, and tagged with *O*-benzyloxyamine hydrochloride (BOA). The BOA-tagged N-glycans are then dissolved in matrix, spotted on a MALDI target plate, and analyzed by MALDI-TOF MS. Using this method, 46 N-glycans were detected in all of the samples, 35 of the detected N-glycans were significantly increased, and 17 of those N-glycans were found to strongly predict and distinguish the BC patients as determined by a ROC AUC of greater than 0.8. Of these 17 N-glycans, only 8 could distinguish the healthy controls from both stage I and II BC patients. Most of the more abundant N-glycans in cancer patients showed more drastic increases in the early stages (I and II) compared to the later stages (III and IV). Almost all glycosylation classes increased in abundance in either the stage I or II BC patients, except for the high-mannose and hybrid N-glycans. While the large differences between the healthy and BC samples is surprising, this study involves a method and subject set never examined for this purpose.

Lee and colleagues used a relatively simple method to analyze N-glycans in serum that entails glycoprotein denaturation with heat, N-glycan release with PNGase F, N-glycan purification with solid-phase extraction (SPE), and mixing

the N-glycans with matrix for MALDI-TOF MS analysis.²²⁰ When comparing 311 healthy volunteers' serum samples with 256 BC patients' serum samples, 30 of the N-glycans were significantly differentially abundant. Performing a ROC analysis incorporating all of these N-glycans resulted in an AUC of 0.93 when comparing the healthy samples to all the BC samples, but if broken down by stages, the AUC increased to 0.96 for differentiating healthy samples from stage I and II BC samples. Interestingly, the AUC when comparing healthy samples to stage III and IV samples decreased to 0.89. Early-stage BC was found to have more significant alterations than late-stage BC in several studies and demonstrates the potential for early BC detection. The only significant alteration found in the glycosylation classes was an increase in high-mannose N-glycans. Both of these studies found many N-glycans to have altered abundances in large BC serum sets by utilizing methods not yet attempted for serum N-glycome-based BC detection.

The findings discussed so far have investigated the total serum N-glycome, but other studies have targeted specific serum glycoproteins for N-glycan profiling and found significant differences in their glycosylation in BC. Abd Hamid and colleagues found that α 1-antichymotrypsin (ACT), haptoglobin β -chain, and α 1-acid glycoprotein (AGP) had N-glycans containing sLe^x epitopes in the BC patient serum but not in the healthy controls serum.²⁰⁷ Increased fucosylation of haptoglobin and AGP N-glycans in BC patients has been found in several studies, while a decrease in fucosylation was seen in transferrin.²²¹⁻²²⁴ Goodarzi and colleagues found α 1-proteinase inhibitor (API) N-glycans had

decreased branching and increased fucosylation in BC patient.²²⁵ Many of these serum glycoproteins have elevated abundances in BC patients.²²⁶

There is a N-linked glycosylation site on the crystallizable fragment (Fc) of immunoglobulin G (IgG) that can drastically alter the ability of the antibody to bind to leukocytes and consequently, effect antibody-dependent cell-mediated cytotoxicity (ADCC).¹²⁷ A multiple logistic regression model incorporating abundances of 7 N-glycans (4 of which were bisected) released from the IgG Fc was able to distinguish healthy controls from BC patients with an AUC of 0.84.²²⁷ Another study found an increase in fucosylated and agalactosylated IgG Fc N-glycans in BC patients.¹⁰² By targeting the carrier proteins of the glycans significantly changing in the total serum, these studies provide an alternative approach for N-glycan profiling-based BC detection that still only requires a minimally invasive and cost-effective blood draw.

A study by Haakensen and colleagues linked serum N-glycan profiling data with transcriptomic analysis to identify associations with specific N-glycan alterations and transcripts representing functional pathways.²²⁸ Using this approach, triantennary trisialylated N-glycans were associated with lower levels of focal adhesion and integrin-mediated cell adhesion involved transcripts, and by integrating tumor characteristics and clinical outcomes, these N-glycans were also associated with poor prognosis in ER negative tumors. While samples with all of these forms of data may be difficult to attain, this type of integrated analysis has the potential to provide the biological context to serum N-glycan biomarkers. This study solely used samples from BC patients, but the promising results

prompt the need for a study comparing BC samples to controls, preferably donated from women with a benign breast disease.

The latest study to perform serum N-glycan profiling of 145 breast cancer patients compared them to 171 healthy females.²²⁹ While three N-glycans (two trisialylated triantennary and one tetrasialylated tetraantennary) were lower in the breast cancer patients, the serum N-glycan profiles of the breast cancer patients differed depending on their molecular subtype and stage. This, like the majority of the serum N-glycan profiling studies discussed here, has contributed to a greater understanding of how N-glycosylation is altered and functions in breast cancer, but it does not address certain clinical areas of need for breast cancer, such as improving the specificity of the screening process.

1.5 Broad Overview

The use of MS across a broad array of disciplines can be partially attributed to the variety of mass spectrometers that utilize different ionization sources and mass analyzers. MALDI is a form of ionization that utilizes chemical matrix to transfer the energy of the MS laser to analyte molecules, even in the presence of contaminants and biological samples with complex compositions. Efforts have been made to improve the sensitivity of MALDI MS analysis for a range of biomolecules by optimizing sample preparation steps.

One important application of MALDI MS has been for the analysis of N-glycans. The composition of enzymatically released N-glycans can be determined using the mass of the ions detected. As N-glycans are non-linear

structures formed without a template like proteins, the well-characterized N-glycan synthesis pathway can be used to infer the structure of the N-glycans. MALDI MS has been used to profile the N-glycans present in tissues, cells, and biofluids. While MALDI MS has a high data acquisition speed compared to other analytical chemistry techniques, especially when paired with a TOF mass analyzer, the extensive sample preparation needed for current N-glycan analysis of biofluids has limited the throughput of these techniques.

High-throughput analysis of N-glycans in clinical samples can be a powerful tool for biomarker discovery because N-glycans can greatly impact the function of their carrier proteins. N-glycosylation has critical roles in nearly every human physiological system and on a molecular level, can impact protein folding, turnover, substrate specificity, and substrate binding efficiency. Biofluids are an attractive sample type for biomarker discovery studies due to their low or non-invasive collection and dynamic compositions responsive to the immune and metabolic status of the individual. In serum and plasma, nearly all of the proteins that make up the bulk of the protein content are N-glycosylated. The combination of functional-relevance and high-prevalence of N-glycans in serum and other biofluids have led to increasing interest in methods capable of large-scale biomarker studies.

One therapeutic area in need of biomarkers is breast cancer. While mammography has become the standard tool for breast cancer screening, a lack of specificity for suspicious findings leads many women to have unnecessary biopsies when they do not have breast cancer. While certain molecular

biomarkers detected in serum have utility in breast cancer subtyping, prognosis, and determining therapeutic sensitivities, there are no effective biomarkers to clarify this clinical issue. Analysis of breast cancer patients' tissues and sera have identified significant alterations in N-glycosylation associated with the presence, molecular subtype, and clinical characteristics of breast cancer tumors, but no study has investigated the serum N-glycan profiles of women that may be impacted by an unnecessary biopsy.

Chapter 2: Hypothesis

2.1 Summary

Breast cancer is responsible for the second highest number of cancer-related deaths annually in US females, and early detection through screening has been shown to save lives.¹⁴⁸ Mammography is the standard screening tool but has issues with false positives and overdiagnosis.¹⁷¹ The identification of a biomarker that could be used as a supplementary tool to inform therapeutic options. The goal of this study is to examine the N-glycan profile of a large cohort of breast cancer patient sera to discover a glycosylation-based biomarker (increase or decrease of the abundance of a N-glycan or glycan class) that can distinguish between cancerous and benign tissue diagnoses. A novel slide-based high-throughput serum N-glycan profiling method will be developed to allow for rapid sample analysis.

The proposed method is adapted from mass spectrometry imaging methods used for tissue N-glycan profiling. Current serum N-glycan analysis methods lack the simplicity and speed found in the novel method, making it a potentially valuable tool for clinical laboratory use. The workflow involves spotting the serum on to a slide, binding the serum proteins to an amine-reactive hydrogel coating, washing off lipids and salts, releasing N-glycans by applying an enzyme, and coating the slide with a matrix. Each spot can be imaged using mass spectrometry to detect a wide range of mass/charge ratios that can be annotated to specific N-glycans. The sample will have an intensity for each N-glycan and these values can be compared across sample groups. By adapting the enzymes

applied to the samples, the method will be able to target specific types of N-glycans.

Once the method is fully developed and validated, it will be applied to a unique clinical cohort that consists of serum collected in a surgical clinic from women with biopsy-confirmed benign or cancerous masses in their breast. We hypothesize that the proposed method will be able to determine potential biomarker targets for breast cancer that could aid in early detection, resulting in less false-positives and unnecessary surgeries.

2.2 Significance

Early detection drastically increases the survival of breast cancer patients and provides multiple treatment options. When diagnosed at its earliest stage, 5-year survival is 98% compared to only 27% for late stage diagnosis.¹⁷⁷ Increased mammography screening has prevented many breast cancer deaths, but it is still limited by current rates of false positives, undetected cancers, overdiagnosis, and overtreatment.²³⁰ Of all women screened by mammography, 12% are recommended to have a follow-up image or biopsy; 95% of those women do not have cancer.¹⁷¹ The combination of false-positives and overdiagnosis can result in unnecessary biopsies, anxiety, and costs. Alternative imaging modalities have not been able to remedy this clinical issue due to being cost-prohibitive or having only modest improvements to specificity.^{175,176}

Given the importance of early detection and widespread application of mammographic screening, there is a need for a more sensitive assay that could

be used to supplement mammography results in order to better inform clinical decision making and limit false positives/overdiagnosis.

This issue of patient screening could be addressed by the detection of a blood-based biomarker. The proposed method is high-throughput and sensitive enough to discover N-glycan biomarkers by analyzing large cohorts of patient samples, but it is also rapid enough to be used as a screening tool to supplement mammography. This assay will not require any changes to standard blood collection protocols. A complicated or uncertain mammogram result could be followed with a serum N-glycan profile to give more confidence as to whether a biopsy is necessary.

2.3 Innovation

The currently used methods for high throughput serum N-glycan analysis are HILIC-UHPLC-FLD, xCGE-LIF, and MALDI-TOF MS.⁸⁷⁻⁹⁰ While these methods have clear utility in understanding the role of serum N-glycans in a variety of diseases, there are elements of each that are incompatible with use in a clinical laboratory.^{91,96,97,231-235} Each of these methods requires labeling or derivatization of released N-glycans and subsequent clean up steps. Not only do these steps involve specialized materials, but they require additional time. To analyze 96 samples with these methods can take from 1 to 5 days.^{89,90,236} The proposed method in Aim 1 can reduce this time by using rapid N-glycan release and avoiding additional steps involved in labelling, derivatization, and clean up. The minimal sample preparation of this method is achieved by anchoring the

glycoproteins directly to the slide. No fractionation or purification, during which sample content can be lost and contaminants could be introduced, is needed with this approach.

The method should be able to be used as a platform to analyze a wide array of biofluids with minimal adaptations to account for different chemical compositions. The integration of automation seems feasible for several steps to increase the precision and throughput even further. Using ion mobility analysis could also aid in detecting structural isomers and provide an additional dimension of information. This method will not only aid in large-scale serum N-glycan studies, but the simplicity, speed, sensitivity, and throughput meet the criteria for clinical laboratory diagnostic tests more than any other method.

2.4 Specific Aim 1

The goal of this aim is to develop an N-glycan profiling method for biofluids that could be used as a clinically diagnostic tool. The current high-throughput serum N-glycan analysis methods have significant sample preparation steps and lengthy detection techniques that make them incompatible with clinical laboratories and limit the potential of N-glycan biomarkers. The workflow involves spotting the serum on an amine-reactive hydrogel coated slide, incubating a binding reaction between the serum proteins and the hydrogel, and washing the slide to remove substances incompatible with mass spectrometry imaging. The N-glycans are enzymatically released and detected using MALDI MS. The method will be evaluated by several criteria including the sensitivity to

low abundance N-glycans, intra- and inter-batch reproducibility, and processing time per sample.

2.5 Specific Aim 2

The goal of this aim is to examine the total serum N-glycan profile of a set of breast cancer patients with MALDI MS and identify cancer-associated N-glycan changes. Alterations in the serum N-glycan profiling have been associated with a variety of cancers and this cohort provides a valuable opportunity to reveal novel N-glycan biomarkers. Specifically, there have been several studies that have found an association between higher branching, fucosylation, and sialylation in the serum of patients with breast cancer.^{102,206,207,216–219} This aim can be accomplished with the method developed in specific aim 1 or with alternative established techniques such as HILIC UHPLC FLD or MALDI MS of HILIC-extracted N-glycans. As the most abundant glycoprotein in blood, IgG contributes a large portion of the N-glycans in serum and there have been several diseases associated with IgG glycosylation changes.¹³³ We can compare the N-glycan profile of IgG and total serum to determine if there are any differences in predictive value or increased sensitivity to N-glycan changes. A patient serum cohort of 298 samples representing benign (n=199) and cancer (n=99) diagnoses is available.

Chapter 3: Serum and Plasma N-glycan Profiling by MALDI- IMS

This chapter has been adapted from a manuscript published in the *Journal of the American Society for Mass Spectrometry*, July 2020 31(12): 2511-2520. CRKB performed all of the experiments and writing of the manuscript. APB, ASM, PMA, and RRD contributed intellectually to the manuscript.

Rapid N-Glycan Profiling of Serum and Plasma by a Novel Slide-Based Imaging Mass Spectrometry Workflow

Calvin R. K. Blaschke, Alyson P. Black, Anand S. Mehta, Peggi M. Angel, and Richard R. Drake

3.2 Introduction

The composition of serum and plasma collected during blood draws reflects dynamic de-cellularized biofluids comprised of thousands of proteins and metabolites. A major component of the fluids are multiple high abundance glycoproteins produced by the liver and immunoglobulins secreted by B-cells. Extensive characterization of the proteomes of serum and plasma has indicated a 10-fold order of magnitude concentration difference from the most abundant proteins, i.e. albumin and immunoglobulins, to the lowest concentration ones.^{72,73} Current proteomic mass spectrometry studies have identified over 4300 individual proteins present in serum,⁷⁴ yet by percentage of abundance, 22 proteins represent 99% of the content.^{72,73} Except for albumin, which is not a glycoprotein and accounts for approximately 50% of the total protein content, the majority of the 22 most abundant proteins and next 30 are glycoproteins secreted by the liver and B-cells. Interrogation of the glycan composition of these most abundant glycoproteins for potential clinical application to differentiate healthy from disease states has been ongoing for decades.⁷⁵⁻⁷⁷ In addition to advanced high-performance liquid chromatography (HPLC)-based methods,²³⁷ in recent years, multiple high-resolution tandem liquid chromatography (LC) and matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) options have been utilized for these types of analyses.^{238,239}

Glycosylation is one of the most common types of post-translational modifications that is estimated to occur on over 50% of proteins.²⁴⁰ The diversity of monosaccharides, inter-saccharide bonding, and non-template derived

branching structures create a large and complex glycan repertoire, called the glycome.²⁴¹ Glycosylation is a highly regulated metabolic process shown to have essential roles in protein folding, molecular trafficking, signal transduction, and many other processes.^{242,243} The most common forms of glycosylation studied for serum or plasma glycoproteins are for those that are O-linked glycans, attached to proteins via threonine/serine residues, and N-glycans, attached to proteins via an asparagine residue. The majority of the most abundant serum and plasma glycoproteins have N-glycosylation. Investigating the changes in specific glycans or glycan structural motifs, such as branching or fucosylation, is a promising area for biomarker identification, and the minimally invasive and cost-effective properties of plasma and serum collection makes the total serum or plasma N-glycome an attractive area of focus.^{75,76} The large cohorts amenable for biomarker discovery and potential clinical applications require high throughput methodologies, that continue to be assessed and evolve.⁹¹

The currently used methods for high throughput serum N-glycome analysis are hydrophilic interaction chromatography (ultra-)high-performance liquid chromatography with fluorescence detection (HILIC-UHPLC-FLD),⁸⁸ multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF),⁸⁹ and matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS).⁹⁰ While these methods have clear utility in understanding the role of serum N-glycans in a variety of diseases, there are elements of each that are incompatible with routine use in a clinical laboratory. Each of these methods requires labeling or derivatization of released

glycans, and further processing and purification steps. Not only do these steps involve specialized materials and expertise, but they require additional time, as analysis of 96 samples with these methods can take from 1 to 5 days.^{89,90,236}

Our collective group has developed methods for analyzing released N-glycans from tissues, cells, and immuno-captured serum glycoproteins by MALDI imaging mass spectrometry (MALDI-IMS).^{19,244–248} Key to each of these methods is the spraying of a molecular coating of peptide N-glycosidase F (PNGase F) onto a slide surface containing the biological sample, followed by detection of released N-glycans by MALDI-IMS. We hypothesized that we could adapt techniques from these workflows to create a rapid and sensitive serum N-glycan profiling method. The sensitivity and repeatability of this method was demonstrated using different control and diseased serum samples, as well as plasma. An additional enzyme, endoglycosidase F3 (Endo F3),²⁴⁹ was used to distinguish core fucose from outer arm fucose locations on the glycan structures. Finally, the utility of this method was demonstrated by analyzing a set of pooled serum samples from patients with either breast cancer or a benign lesion in their breast. The clinical applicability of this serum and plasma N-glycan profiling method improves the potential of glycans to be used as biomarkers. The speed and simplicity of the novel N-glycan profiling method described in this study are compatible with the requirements of clinical laboratories.

3.3 Methods

3.3.1 Materials

Hydrogel coated slides (Nexterion® Slide H) were obtained from Applied Microarrays (Tempe, AZ). The rotary tool was a Dremel 200 series. The well slide module (ProPlate Multi-Array Slide System, 64-well) was obtained from Grace Bio-Laboratories (Bend, OR). Sodium bicarbonate, trifluoroacetic acid, and α -cyano-4-hydroxycinnamic acid were obtained from Sigma-Aldrich (St. Louis, MO). HPLC grade water, acetonitrile, glacial acetic acid, and chloroform were obtained from Fisher Scientific (Hampton, NH). Ethanol was obtained from Decon Labs (King of Prussia, PA). Peptide-N-glycosidase F (PNGase F) PRIME™ was from N-Zyme Scientifics (Doylestown, PA). Endoglycosidase F3 (Endo F3) and sialidase was cloned, expressed, and purified in-house as previously described.²⁴⁹

3.3.2 Samples

A pooled human serum sample was used that represented over 360 healthy human donors.²⁵⁰ A control plasma-EDTA sample was obtained from a healthy donor. Serum from non-obese or obese patients with benign breast lesions or breast cancer was pooled from 5 - 10 patients, derived from a previously published study.²⁵¹

3.3.3 Array Preparation

Hydrogel coated slides were used, and wells were created by attaching a 64 (4 x 16) well module. The hydrogel coated slides were equilibrated to room temperature in a moisture resistant pouch for 45 minutes. The slide was ground down with a rotary tool until it fit into a Bruker MTP Slide Adapter II. The well module was attached to the slide, the wells were outlined with a marker on the uncoated side of the slide, and then the well module was unattached.

3.3.4 Sample Capture and Washing

A μl of serum was diluted in 2 μl of sodium bicarbonate (100 mM, pH 8.0). After briefly mixing with the pipette, 1 μl was spotted within the outline of a well. Only wells from the two innermost rows were used. The pooled serum standard and plasma standard were spotted in quadruplicate, and the clinical serum samples were spotted in triplicate. Spots were left to immobilize to the slide at room temperature for 1 hr in a humidity chamber made from a culture dish with a Wypall X 60 paper towel lining the bottom and two rolled KimWipes saturated with distilled water on opposite sides. For the immobilization, the slide was placed flat in the dish in between the two KimWipes. The slide was then dried in a desiccator for 15 minutes and the well module was attached. Each well was filled with 50 μl of Carnoy's solution (10% glacial acetic acid, 30% chloroform, 60% 200 proof ethanol) without mixing or agitating for 3 minutes. The Carnoy's solution was dumped out of the well and the wash was repeated 2 more times. After the 3 Carnoy's solution washes, each well was rinsed with 50 μl of double

distilled water without mixing or agitating for 1 minute. Following removal of water, the slide with the well module attached was dried in a desiccator for 30 minutes.

3.3.5 N-Glycan Release and Matrix Application

To release the N-glycans from the immobilized proteins, the well module was removed and a M5 TM-Sprayer (HTX Technologies) was used to spray a 0.1 mg/ml PNGase F PRIME solution in water with 15 passes at 25 μ l/min, 1200 mm/min, 45°C, 3 mm spacing between passes with 10 psi nitrogen gas. The slide was incubated for 2 hours at 37°C in a preheated humidity chamber. To apply the MALDI matrix α -cyano-4-hydroxycinnamic acid (CHCA, 7 mg/mL in 50% acetonitrile/0.1% trifluoroacetic acid), the same automated sprayer was used with 10 passes at 100 μ l/min, 1200 mm/min, 79°C, 2.5 mm spacing between passes with 10 psi nitrogen gas. For Endo F3 or sialidase treatment, a mix of 10% Endo F3 or sialidase and 90% PNGase F Prime was sprayed at the same settings as described for only PNGase F and all other steps remained the same.^{247,249}

3.3.6 MALDI imaging mass spectrometry

A SolariX Legacy 7T Fourier-Transform Ion Cyclotron Resonance (FTICR) mass spectrometer (Bruker) equipped with a MALDI source was used to image the slides. Images were collected with a SmartBeam II laser operating at 2000 Hz with a 25 μ m laser spot size and a smartwalk pattern at a 250 μ m raster with

200 laser shots per pixel. Samples were analyzed in positive ion, broadband mode using a 512k word time domain spanning a m/z range of 500-5000. Individual regions to be imaged were manually outlined to encompass each well with a sample, as well as one blank well that did not have any sample spotted in it. An on-slide resolving power of 48,000 at $m/z = 1136$ was calculated.

3.3.7 Data Analysis

Data was imported at a 0.95 ICR reduction noise threshold and normalized to total ion current in to FlexImaging v5.0 (Bruker) for manual N-glycan peak selection based on theoretical mass values. Data Analysis 5.0 (Bruker) was used for spectra recalibration with a linear function based on N-glycan theoretical masses. Spectra were imported into SCiLS Lab software 2017a (Bruker) for individual peak visualization and quantification. Each well was designated a unique region and area under monoisotopic peak values were exported from each well. Individual N-glycan area under monoisotopic peak values were divided by the sum of the area under monoisotopic peak values of all detected N-glycans for relative intensity values. Quantifications of N-glycan structural classes were calculated by summing the relative intensities of the individual N-glycans belonging to each class, as determined by the putative structures. Each slide contained non-sample wells processed in the same manner as wells with samples, and can be used for background signal subtraction in subsequent data analysis steps.

3.3.8 MS/MS

A timsTOF fleX mass spectrometer (Bruker) was used to acquire MS/MS data by collision-induced dissociation (CID) for N-glycans directly from the processed slides. A 2 Da window was used for MS/MS precursor selection. The collision energies were individually optimized for consistent and sensitive fragmentation and ranged from 100-140 eV. The number of laser shots summed were also optimized per N-glycan for fragmentation reporting. Detected glycan species were cross-referenced with existing in-house structural databases from previous MALDI-FTICR-MS studies.^{17,203,249}

3.4 Results

This study was initiated to identify a serum and plasma glycan profiling strategy based on direct spotting of sample on glass slides, spraying of PNGase F to release N-glycans, followed by detection using MALDI-IMS. The goal was to develop a protocol that was both rapid (relative to current methods used) and reproducible. Analysis workflows for processing tissues on slides for N-glycan MALDI-IMS analysis using sprayed PNGase F to release N-glycans are established and used in this study.²⁴⁷ It was the initial sample spotting and processing steps prior to PNGase F spraying that required extensive evaluation and optimization. Parameters like slide chemistry, washing and sample buffers, sample amount, and digestion conditions were assessed. Peak intensities and numbers of N-glycans detected by MALDI-FTICR-MS were the final determinants.

Initially, multiple slide chemistries were evaluated for serum spotting, including blank glass histology slides and different modifications like indium tin oxide (ITO), poly-lysine and nitrocellulose. Because MALDI was being used for detection, a desalting component was needed. A denaturing component was also required in the sample preparation due to the high concentration of protein present in even 1 μ l of human serum (~75 μ g) and the complexities of releasing N-glycans from native proteins. Each of the indicated slide chemistries had issues with sample loss during the washing and denaturing steps leading to inconsistent or loss of N-glycan detection across replicates (data not shown). Attempts to fixate the samples to the slides using formalin cross-linking resulted in sample spot hardening and flaking off of the slide after washing. To address these limitations, amine reactive slides were assessed which provided covalent attachment of the serum proteins, facilitating efficient desalting and denaturation steps. These slides were used throughout the study as summarized in the workflow schematic in shown in Figure 6.

The amine reactive hydrogel slides used contain N-hydroxysuccinimide (NHS) esters free to covalently react with primary amines. Serum samples were mixed with 100 mM sodium bicarbonate at pH 8 to dilute the sample for more efficient capture on the hydrogel coated slide. After 1 hour incubation, the bound proteins on the slide can be washed with minimal loss of sample. Carnoy's solution (10% glacial acetic acid, 30% chloroform, 60% ethanol) was used to delipidate and denature the proteins, and water washes were done to remove residual Carnoy's solution and salts. PNGase F was sprayed evenly over the

slide, and after a 2 hour incubation, a CHCA matrix coating was sprayed on the slide. The slide was analyzed using MALDI-FTICR MS, resulting in an overall mass spectrum that can be used to obtain heat map images correlating to the intensities of specific m/z peaks across the slide. The intensities of the detected N-glycans can be quantified for areas covering the wells of specific samples, allowing for comparisons of N-glycans between samples.

The current workflow described here (Figure 6) will analyze 28 spots per slide and additional slides can be prepared together up to the point of the MALDI IMS with little additional time. The total time to prepare a slide for imaging is 6 hrs. Each spot can be imaged in 10 minutes with the given parameters. The entire workflow can be used to analyze a single spot in less than 6.5 hours and 28 spots in less than 11 hrs. To define the breadth of N-glycans detected, a serum and plasma standard was spotted in quadruplicate, processed and analyzed. Using this method, 75 N-glycans were detected in human serum (Supplemental Table 1) with an average coefficient of variation (CV) of 6% for the 20 most abundant glycans.

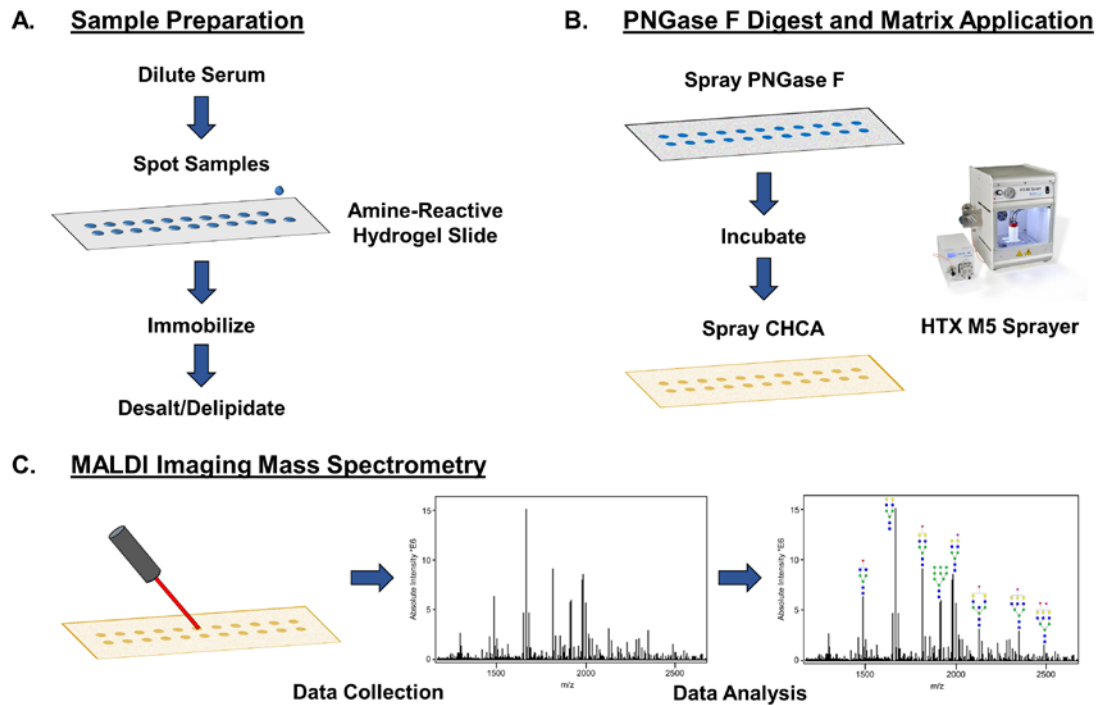


Figure 6. Workflow for slide-based serum and plasma N-glycan profiling by MALDI IMS. (A) Sample preparation steps include diluting the serum or plasma 1:2 in sodium bicarbonate (100 mM, pH 8.0) and spotting 1 microliter on to a hydrogel-coated slide. The sample spots are put in a humidity chamber for an hour to immobilize to the slide. Carnoy's solution is used to delipidate the samples for sensitive N-glycan detection. (B) The sample spots are sprayed with PNGase F to enzymatically release N-glycans, and CHCA matrix is applied. (C) Sample spots are imaged in distinct regions by a MALDI-FTICR-MS, allowing for images and abundances of m/z peaks corresponding to N-glycans to be compared across samples.

Observed Mass	Error (ppm)	Composition	Glycosylation Traits	Observed Mass	Error (ppm)	Composition	Glycosylation Traits
1136.3936	-2.46	Hex3HexNAc3 + 1Na		2179.7617	7.2	Hex5HexNAc5NeuAc1 + 2Na	Tri, Sia
1257.4178	-3.82	Hex5HexNAc2 + 1Na	Man	2245.7654	-3.03	Hex5HexNAc4NeuAc2 + 1Na	Bi, Sia
1282.4618	5.85	Hex3dHex1HexNAc3 + 1Na	Fuc	2246.7704	-9.88	Hex5dHex2HexNAc4NeuAc1 + 1Na	Fuc, Bi, Sia
1298.4501	0.69	Hex4HexNAc3 + 1Na		2267.7636	0.71	Hex5HexNAc4NeuAc2 + 2Na	Bi, Sia
1339.471	-3.51	Hex3HexNAc4 + 1Na	Bi	2268.7652	-7.57	Hex5dHex2HexNAc4NeuAc1 + 2Na	Fuc, Bi, Sia
1419.4712	-3.03	Hex6HexNAc2 + 1Na	Man	2289.7561	10.26	Hex5HexNAc4NeuAc2 + 3Na	Bi, Sia
1444.5052	-1.32	Hex4dHex1HexNAc3 + 1Na	Fuc	2303.843	12.54	Hex5dHex1HexNAc5NeuAc1 + 1Na	Tri, Fuc, Sia
1460.5025	0.34	Hex5HexNAc3 + 1Na		2304.8372	1.17	Hex5dHex3HexNAc5 + 1Na	Tri, Fuc
1485.5323	-0.94	Hex3dHex1HexNAc4 + 1Na	Fuc, Bi	2319.828	8.19	Hex6HexNAc5NeuAc1 + 1Na	Tri, Sia
1501.5287	0.07	Hex4HexNAc4 + 1Na	Bi	2320.8293	-0.04	Hex6dHex2HexNAc5 + 1Na	Tri, Fuc
1542.5362*	-12.25	Hex3HexNAc5 + 1Na	Bis	2325.806	0.9	Hex5dHex1HexNAc5NeuAc1 + 2Na	Tri, Sia, Fuc
1581.5275	-0.44	Hex7HexNAc2 + 1Na	Man	2341.7948	-1.71	Hex6HexNAc5NeuAc1 + 2Na	Tri, Sia
1606.5643	2.74	Hex5dHex1HexNAc3 + 1Na	Fuc	2391.8211	-3.76	Hex5dHex1HexNAc4NeuAc2 + 1Na	Fuc, Bi, Sia
1611.5181	-5.27	Hex4HexNAc3NeuAc1 + 2Na	Sia	2393.8442	-0.67	Hex7HexNAc6 + 1Na	Tetra
1647.5813	-3.16	Hex4dHex1HexNAc4 + 1Na	Fuc, Bi	2413.8225	1.08	Hex5dHex1HexNAc4NeuAc2 + 2Na	Fuc, Bi, Sia
1663.5813	-0.06	Hex5HexNAc4 + 1Na	Bi	2435.8429	13.63	Hex5dHex1HexNAc4NeuAc2 + 3Na	Fuc, Bi, Sia
1688.6115*	-0.89	Hex3dHex1HexNAc5 + 1Na	Fuc, Bis	2449.8932	8.65	Hex5dHex2HexNAc5NeuAc1 + 1Na	Tri, Fuc, Sia
1704.6006	-4.28	Hex4HexNAc5 + 1Na	Bis	2465.8986	12.86	Hex6dHex1HexNAc5NeuAc1 + 1Na	Fuc, Tri, Sia
1743.581	0	Hex8HexNAc2 + 1Na	Man	2466.9006	5.39	Hex6dHex3HexNAc5 + 1Na	Fuc, Tri
1751.573	-13.93	Hex5HexNAc3NeuAc1 + 1Na	Sia	2487.8561	-0.24	Hex6dHex1HexNAc5NeuAc1 + 2Na	Fuc, Tri, Sia
1773.571	-9.13	Hex5HexNAc3NeuAc1 + 2Na	Sia	2610.9273	8.77	Hex6HexNAc5NeuAc2 + 1Na	Tri, Sia
1792.6277	2.06	Hex4HexNAc4NeuAc1 + 1Na	Sia, Bi	2611.9374	4.82	Hex6dHex2HexNAc5NeuAc1 + 1Na	Fuc, Tri, Sia
1809.6462	3.81	Hex5dHex1HexNAc4 + 1Na	Fuc, Bi	2632.9301	13.65	Hex6HexNAc5NeuAc2 + 2Na	Tri, Sia
1850.6678	1.03	Hex4dHex1HexNAc5 + 1Na	Fuc, Bis	2633.9313	6.34	Hex6dHex2HexNAc5NeuAc1 + 2Na	Fuc, Tri, Sia
1866.6728	6.43	Hex5HexNAc5 + 1Na	Bis	2638.9272	14.44	Hex5dHex1HexNAc5NeuAc2 + 3Na	Fuc, Tri, Sia
1905.6358	1.05	Hex9HexNAc2 + 1Na	Man	2684.9569	5.85	Hex7HexNAc6NeuAc1 + 1Na	Tetra, Sia
1938.6922	5.31	Hex4dHex1HexNAc4NeuAc1 + 1Na	Fuc, Sia, Bi	2685.9739	4.58	Hex7dHex2HexNAc6 + 1Na	Tetra, Fuc
1954.6722	-2.35	Hex5HexNAc4NeuAc1 + 1Na	Bi, Sia	2706.9205	-3.88	Hex7HexNAc6NeuAc1 + 2Na	Tetra, Sia
1955.6731	-12.32	Hex5dHex2HexNAc4 + 1Na	Fuc, Bi	2757.9599	-8.27	Hex6dHex3HexNAc5NeuAc1 + 1Na	Tri, Fuc, Sia
1960.6604	-1.73	Hex4dHex1HexNAc4NeuAc1 + 2Na	Fuc, Sia, Bi	2779.9401	-11.65	Hex6dHex3HexNAc5NeuAc1 + 2Na	Tri, Fuc, Sia
1976.6556	-5.56	Hex5HexNAc4NeuAc1 + 2Na	Sia, Bi	2902.0267	9.27	Hex6HexNAc5NeuAc3 + 1Na	Tri, Sia
2012.7304	5.81	Hex5dHex1HexNAc5 + 1Na	Fuc, Bi	2923.9973	2.63	Hex6HexNAc5NeuAc3 + 2Na	Tri, Sia
2028.7138	0.1	Hex6HexNAc5 + 1Na	Tri	2976.0366*	-14.85	Hex7HexNAc6NeuAc2 + 1Na	Tetra, Sia
2100.7473	6	Hex5dHex1HexNAc4NeuAc1 + 1Na	Fuc, Bi, Sia	2977.0911	11.45	Hex7dHex2HexNAc6NeuAc1 + 1Na	Tetra, Fuc, Sia
2101.7489	-2.95	Hex5dHex3HexNAc4 + 1Na	Fuc, Bi	3049.0833	1.71	Hex6dHex3HexNAc5NeuAc2 + 1Na	Tri, Fuc, Sia
2122.715	-4.48	Hex5dHex1HexNAc4NeuAc1 + 2Na	Fuc, Sia, Bi	3050.0908	5.71	Hex8HexNAc7NeuAc1 + 1Na	Tetra, Sia
2158.7939	8.01	Hex5dHex2HexNAc5 + 1Na	Fuc, Tri	3071.0456	-7.25	Hex6dHex3HexNAc5NeuAc2 + 2Na	Tri, Fuc, Sia
2174.7955	11.04	Hex6dHex1HexNAc5 + 1Na	Fuc, Tri				

Table 1. N-glycans observed in serum and plasma. N-glycans marked with an asterick (*) were only seen in serum samples from patients with benign lesions or breast cancer. Man = high mannose, Fuc = fucosylated, Com = complex, Bi = bi-antennary, Tri = tri-antennary, Tetra = tetra-antennary, Bis = bisected, Sia = sialylated.

For sialylated N-glycans, predominant peaks detected in formalin-fixed paraffin embedded (FFPE) tissues have an extra Na ion per each sialic acid present.²⁴⁵ This serum and plasma protocol results in detection of many single sodiated sialylated glycans, in addition to the extra Na adduct glycoforms. This is also seen in frozen tissue analyses of N-glycans,²⁴⁶ and is likely related to the Carnoy's and water washes not done with FFPE tissues. An equivalent number of N-glycans were detected in plasma, and the CVs for top 20 most abundant N-glycans and structural classes were 9% and 5%, respectively. The average mass error for all detected glycans was 0.01 ± 0.01 Da. A representative spectrum obtained from the serum standard and plasma standard is shown in Figure 7.

To evaluate the repeatability and day-to-day variation of this workflow, a human serum standard and human plasma standard was spotted and processed using the optimized workflow, and the workflow was repeated on the next two successive days (day 1-3). Figure 8 displays the area of the 20 most abundant glycan compositions in serum using area under the peak intensities compared to the total sum of peak intensity values for all of the detected glycan peaks. The most abundant glycan peaks included species across a broad m/z range ($m/z = 1136$ to $m/z = 3071$). For sialylated glycan species that had multiple sodiated adducts detected, only one species was included in Figure 8. These glycans showed little to moderate variability across the experiments, with the CV averaging to 11.8%, with some peaks having much lower variation.

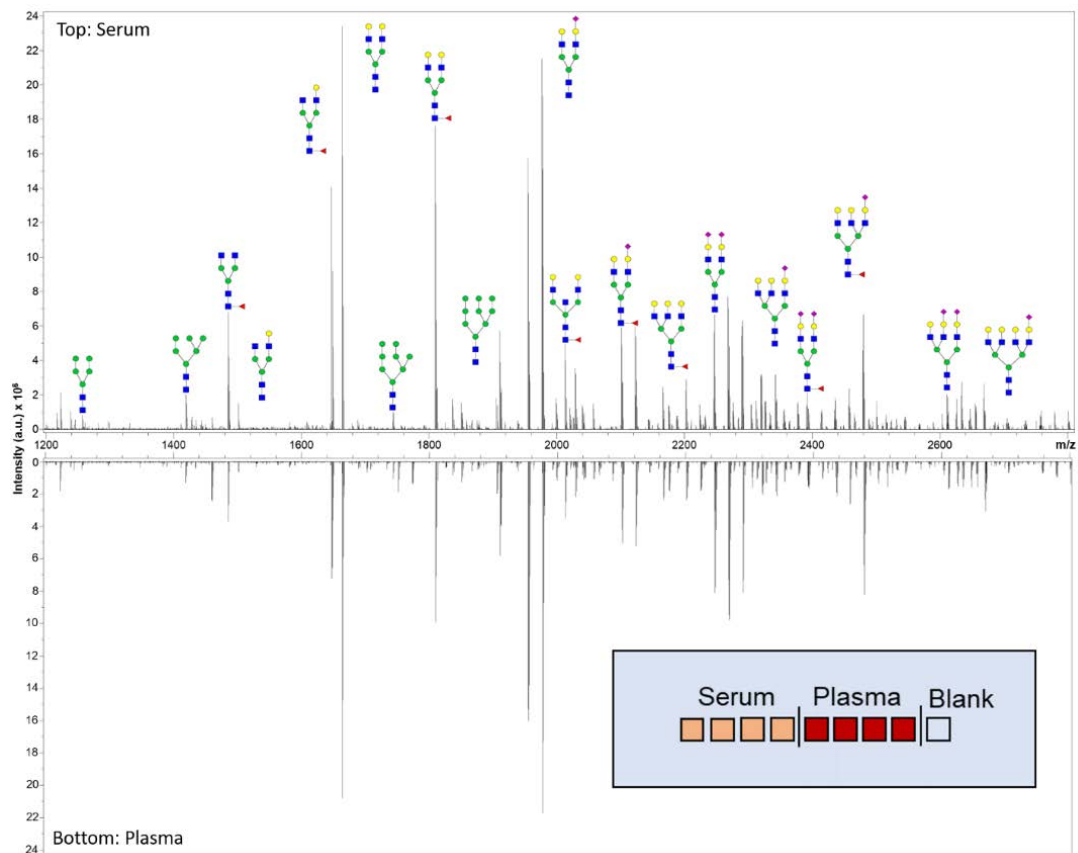


Figure 7. Representative spectra of a standard serum sample (top) and a standard plasma sample (bottom) analyzed with the N-glycan profiling method by MALDI-FTICR MS. A sample of detected N-glycans are annotated with putative structures. The experimental layout of the slide analyzed is displayed in the bottom right.

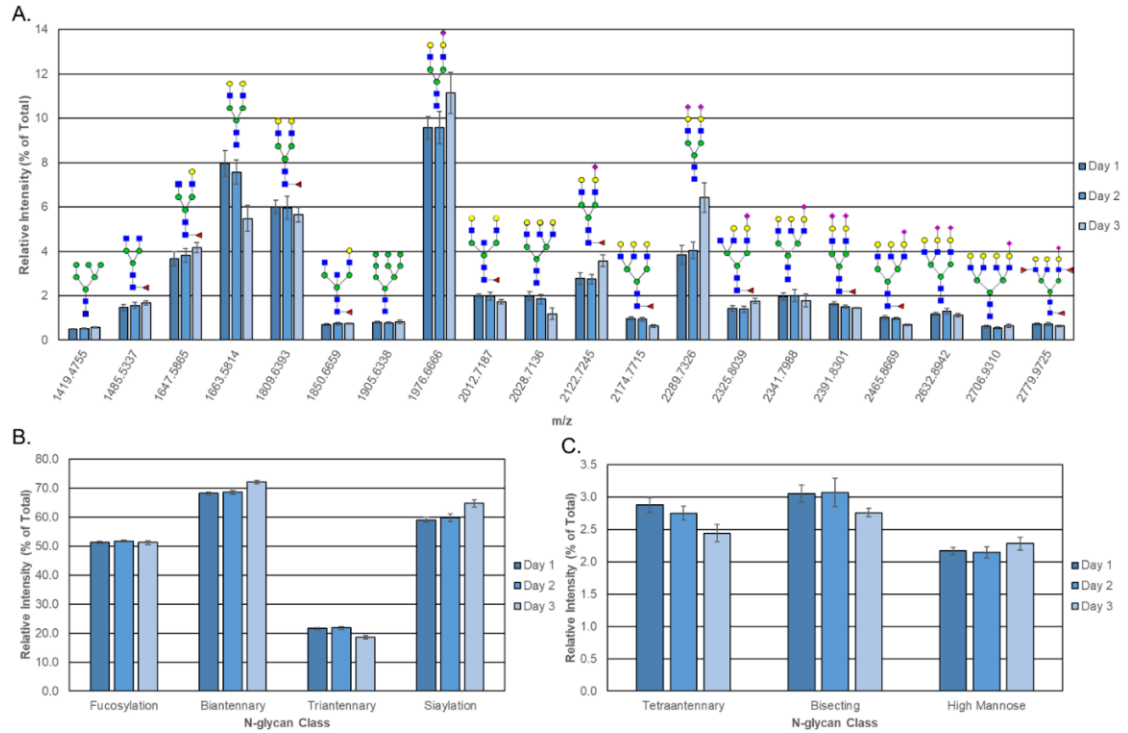


Figure 8. Repeatability of the N-glycan profiling method for serum analysis. The same serum standard sample was analyzed by MALDI-FTICR MS on three successive days with a blank well that contained no serum on each run. (A) The graph shows the average relative intensities detected for the twenty most abundant N-glycans (normalized to the overall sum of N-glycan intensities) with putative structures. For determining the most abundant N-glycans, only N-glycans with no overlap with another N-glycan peak isotope were included, and only one species was selected for sialylated N-glycans with multiple sodiated adducts. Structural classes were assigned to the N-glycans, and the sum of the N-glycans for the (B) high abundance and (C) low abundance classes are displayed. Error bars indicate the standard deviation of the four replicates.

If the glycans were grouped based on structural features, this improved repeatability across the three days (CV = 5%). This type of grouping and improvement in variability has been previously reported for MALDI serum glycan analysis.⁹¹ The average relative intensity of the most abundant unique N-glycan peak ($m/z = 1976.6666$, Hex5HexNAc4NeuAc1 + 2Na) across all experiments was 10% (SD \pm 1%).

The results of the plasma standard showed larger variability within the repeatability parameters (16% and 8% CV for the twenty most abundant unique N-glycan peaks and N-glycan structural classes, respectively) (Figure 9). Further optimization studies for plasma and the different anti-coagulants (EDTA, heparin, citrate) used clinically are still ongoing.

A mixture of Endo F3 or sialidase and PNGase F was applied to the serum and plasma samples. An overview of the enzymatic treatments is shown in Figure 10. Endo F3 is a glycosidase that cleaves N-glycans between the first two N-acetylglucosamines (GlcNAc) linked to asparagine, preferentially when the first GlcNAc is fucosylated. It can be used to differentiate N-glycans containing the core fucose versus glycans with fucose modifications on the antennae.²⁴⁹ Additionally, the m/z 's of Endo F3 cleaved N-glycans products can be calculated based on the loss of 349 m.u. and confirmed by MS/MS. Use of the Endo F3/PNGase F mix allowed comparison of the data already obtained for fucosylated N-glycans using PNGase F alone. This modification requires no increase in processing time and no changes to the MALDI IMS parameters.

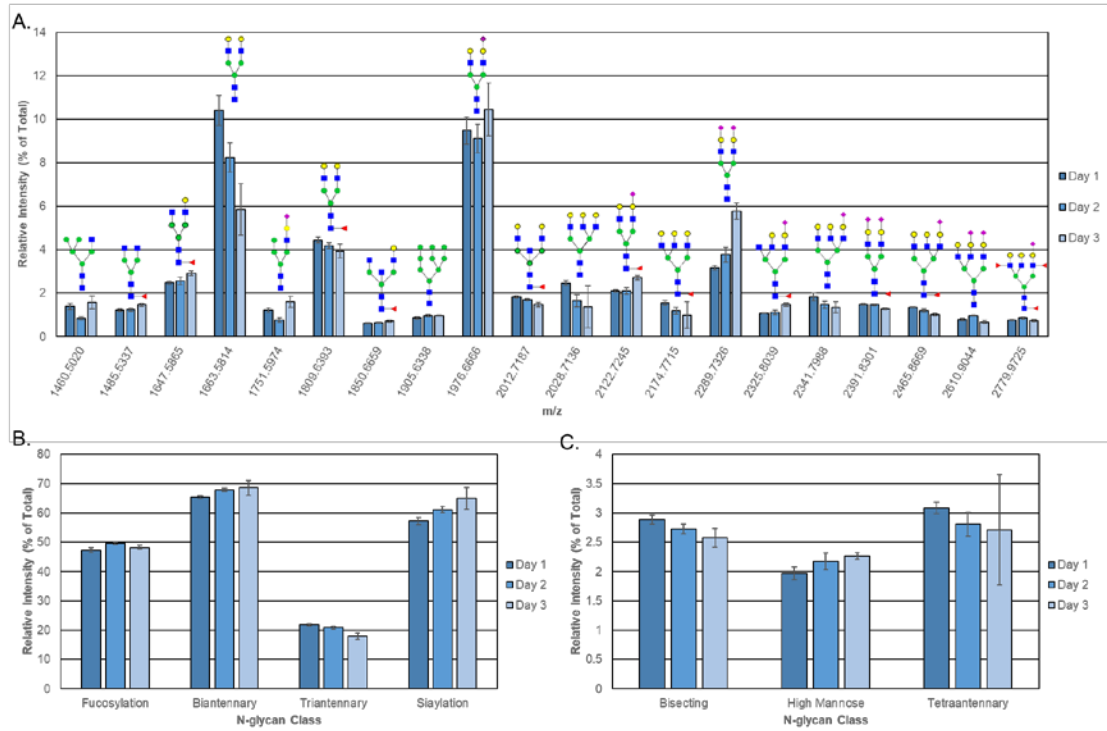


Figure 9. Repeatability of the N-glycan profiling method for plasma analysis. The same plasma standard sample was analyzed on three successive days with a blank well that contained no serum on each run. (A) The graph shows the average relative intensities detected for the twenty most abundant N-glycans (normalized to the overall sum of N-glycan intensities) with putative structures. For determining the most abundant N-glycans, only N-glycans with no overlap with another N-glycan peak isotope were included, and only one species was selected for sialylated N-glycans with multiple sodiated adducts. Structural classes were assigned to the N-glycans, and the sum of the N-glycans for the (B) high abundance and (C) low abundance classes are displayed. Error bars indicate the standard deviation of the four replicates.

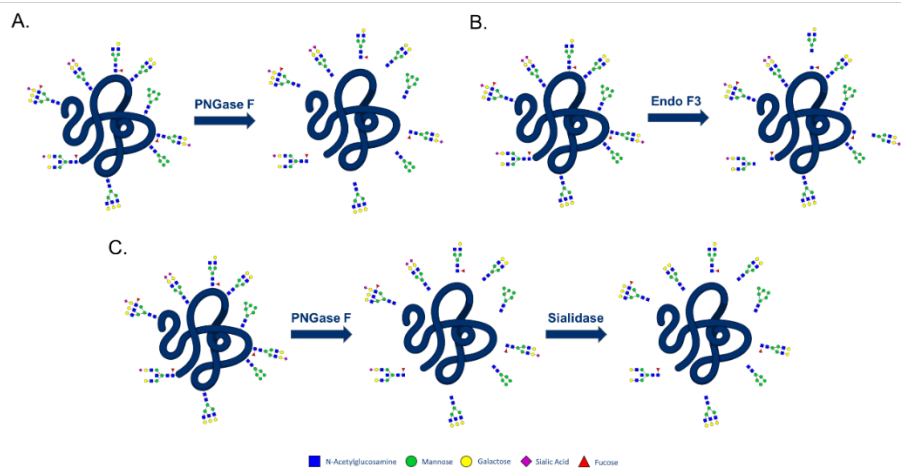


Figure 10. Overview of the enzymatic treatments applied for serum analysis.

As shown in Figure 11 A, multiple fucosylated N-glycans from serum were decreased in abundance in the Endo F3/PNGase F treatment compared to the PNGase F only treatment. Endo F3/PNGase F treatment did not affect the abundances of non-fucosylated glycans in serum and plasma (Figure 11 B). While some N-glycans could be readily identified as core fucosylated using this method (Figure 11 C-D), the identification of several other N-glycan fucosylation locations were obscured by several different factors. Various Endo F3 products have the m/z of a fucosylated N-glycan, making it difficult to attribute the signal intensity to either N-glycan. Other Endo F3 products that did not have an overlapping m/z with another N-glycan were not detected, but their parent N-glycan abundance did decrease. Collision-induced dissociation MS/MS was also used to ascertain structural information, with representative data examples provided in Figure 12.

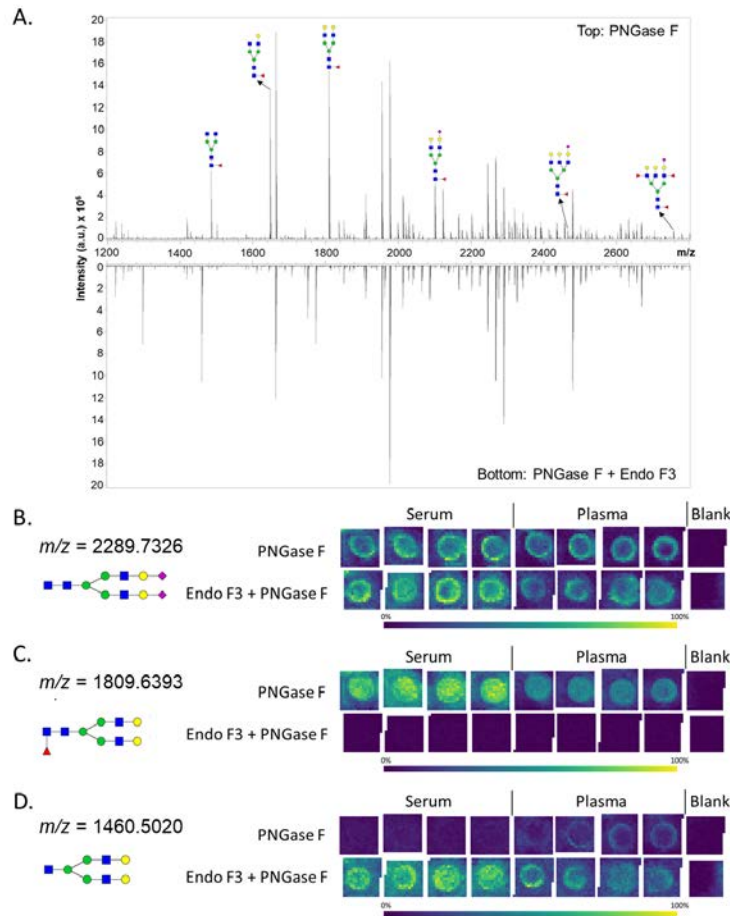


Figure 11. Endo F3 modification of the N-glycan profiling method for serum analysis. (A) Representative spectrum for a serum standard treated with PNGase F (top) or PNGase F and Endo F3 (bottom) with annotations for a set of putative N-glycan structures. The N-glycan profile of a serum standard and a plasma standard was analyzed by MALDI-FTICR MS in quadruplicate by a treatment with PNGase F or a PNGase F and Endo F3 mix. The intensities of a (B) non-core fucosylated N-glycan (m/z 2289.7326), (C) core-fucosylated N-glycan (m/z 1809.6393), and (D) Endo F3 product (m/z 1460.5020), respectively, are displayed for both treatments.

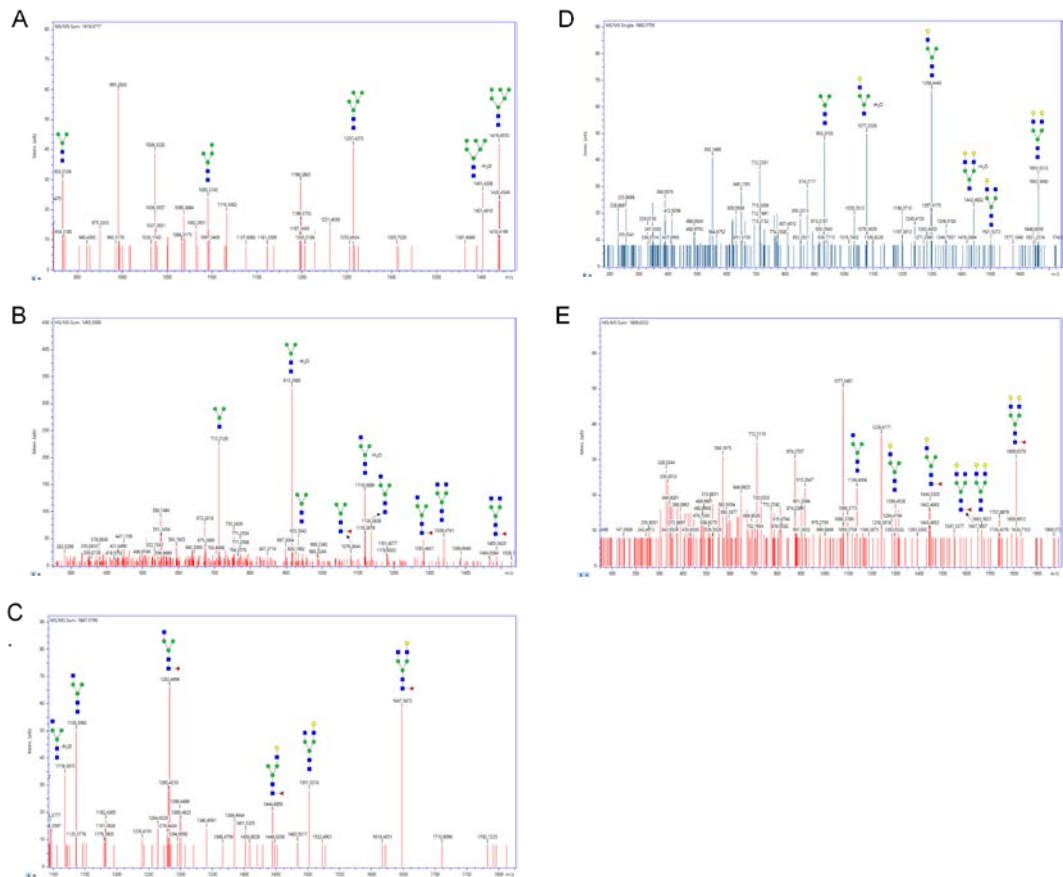


Figure 12. MS/MS of serum N-glycans by CID done directly on the slide. A) $m/z = 1419.4755$, B) $m/z = 1485.5337$, C) $m/z = 1647.5865$, D) $m/z = 1663.5814$, and E) $m/z = 1809.6393$ N-glycan structures were confirmed with MALDI-TOF MS/MS. The collision energies and number of laser shots were individually optimized for each N-glycan.

An alternative enzymatic treatment of sialidase and PNGase F were concurrently applied to the samples as well. Sialidase is an endoglycosidase that releases sialic acid residues from the terminal end of the antennas on N-glycans. This treatment collapses the intensity of all the differentially sialylated structures and their alternatively sodiated structures of an N-glycan into one peak (Figure 13). In this way, the intensity of the non-sialylated branched structures can be analyzed.

The use of these alternative enzymatic treatments increased the sensitivity of the workflow for the detection of core fucosylated and polylactosamines for Endo F3 and sialidase, respectively (Figure 14). Sialidase and PNGase F treatment increases the intensity of polylactosamine N-glycans when compared to only PNGase F treatment of the same sample (Figure 14 A). The intensities of some core fucosylated N-glycans from a sample treated with Endo F3 only were higher than when the sample was treated with PNGase F only.

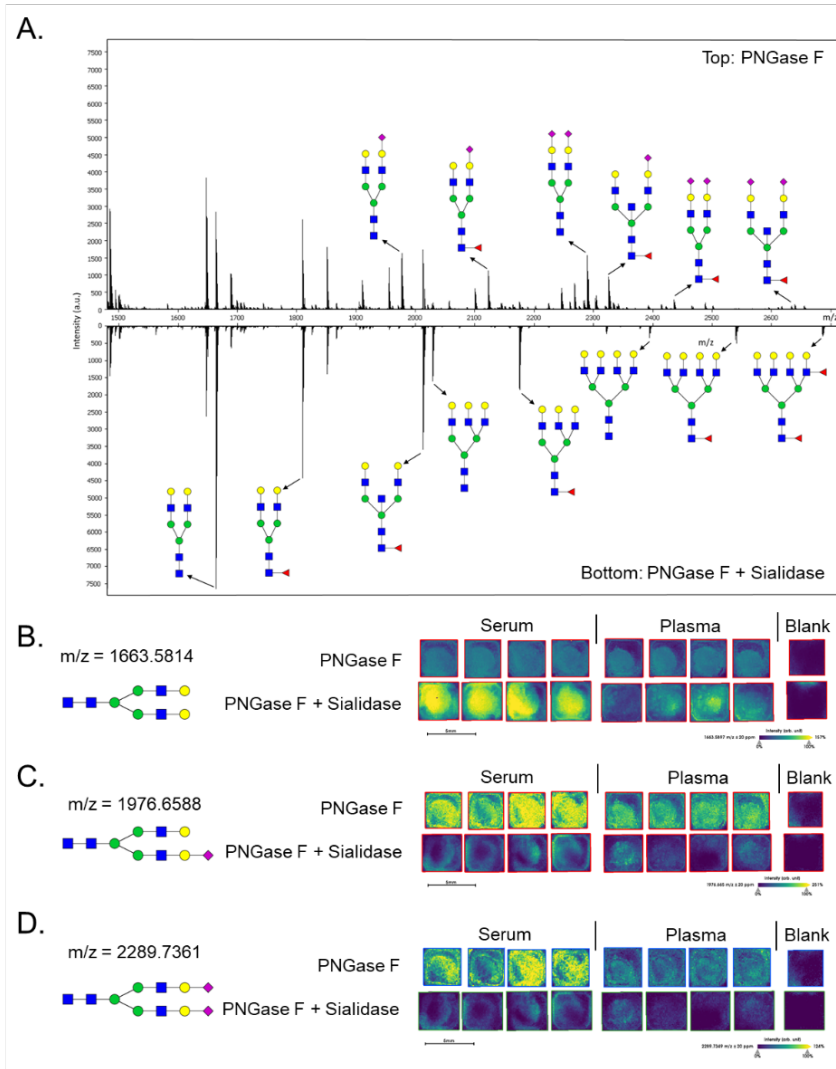


Figure 13. Sialidase modification to the workflow. (A) Representative spectrum for a serum standard treated with PNGase F (top) or PNGase F and Sialidase (bottom) with annotations for a set of putative N-glycan structures. The N-glycan profile of a serum standard and a plasma standard was analyzed. The intensities of a (B) non-sialylated N-glycan (m/z 1663.5814), (C) singly sialylated N-glycan (m/z 1976.6588), and (D) double sialylated N-glycan (m/z 2289.7361), respectively, are displayed for both treatments.

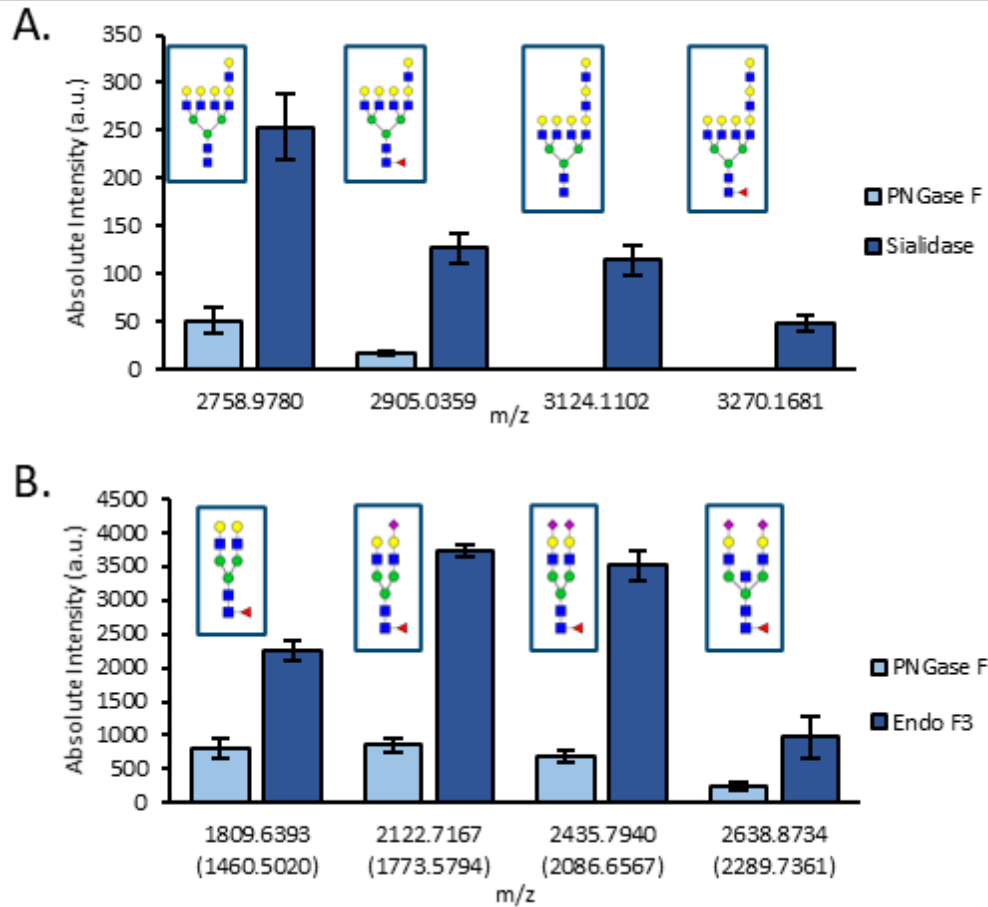


Figure 14. Alternative enzymatic treatments increase sensitivity to classes of N-glycans. (A) Intensities of polylactosamine N-glycans of a sample treated with either PNGase F or PNGase F and sialidase. (B) Intensities of core fucosylated N-glycans of a sample treated with either PNGase F or Endo F3. The m/z of the Endo F3 products that correlate to the PNGase F products are shown in parenthesis.

The potential clinical application of this method was assessed by analyzing a small set of pooled serum samples from non-obese or obese patients with benign breast lesions or breast cancer.²⁵¹ Most of the N-glycans that were detected in the pooled serum standard were also present in these samples, and there were N-glycans that were only found in the samples and not in the serum standard (marked with an asterisk in Table 1). While major differences were not detected across groups for most of the detected N-glycans, some representative examples of differential detection per clinical group are shown in Figure 15. A higher abundance in a core fucosylated, non-galactosylated N-glycan at $m/z = 1485.5337$ (Hex3dHex1HexNAc4 + 1Na) was observed in the cancerous obese serum (Figure 15 A). A mono-galactosylated N-glycan $m/z = 1501.5286$ (Hex4HexNAc4 + 1Na) had a higher abundance in both non-obese samples (Figure 15 B). The non-obese and obese benign samples also had higher amounts of the highly abundant fucosylated bi-antennary N-glycan $m/z = 1809.6393$ (Hex5dHex1HexNAc4 + 1Na) (Figure 15 C). The benign obese serum had a lower abundance of the presumably bisected N-glycan $m/z = 1866.6608$ (Hex5HexNAc5 + 1Na) (Figure 15 D). These N-glycans vary largely in their average relative intensities with $m/z = 1809.6393$ accounting for 5.1% of the glycan signal and $m/z = 1866.6608$ accounting for 0.4%. The N-glycans detected in the clinical serum set had a similar level of variability as those found in the serum standards. The top 20 most abundant N-glycans averaged a CV of 8%.

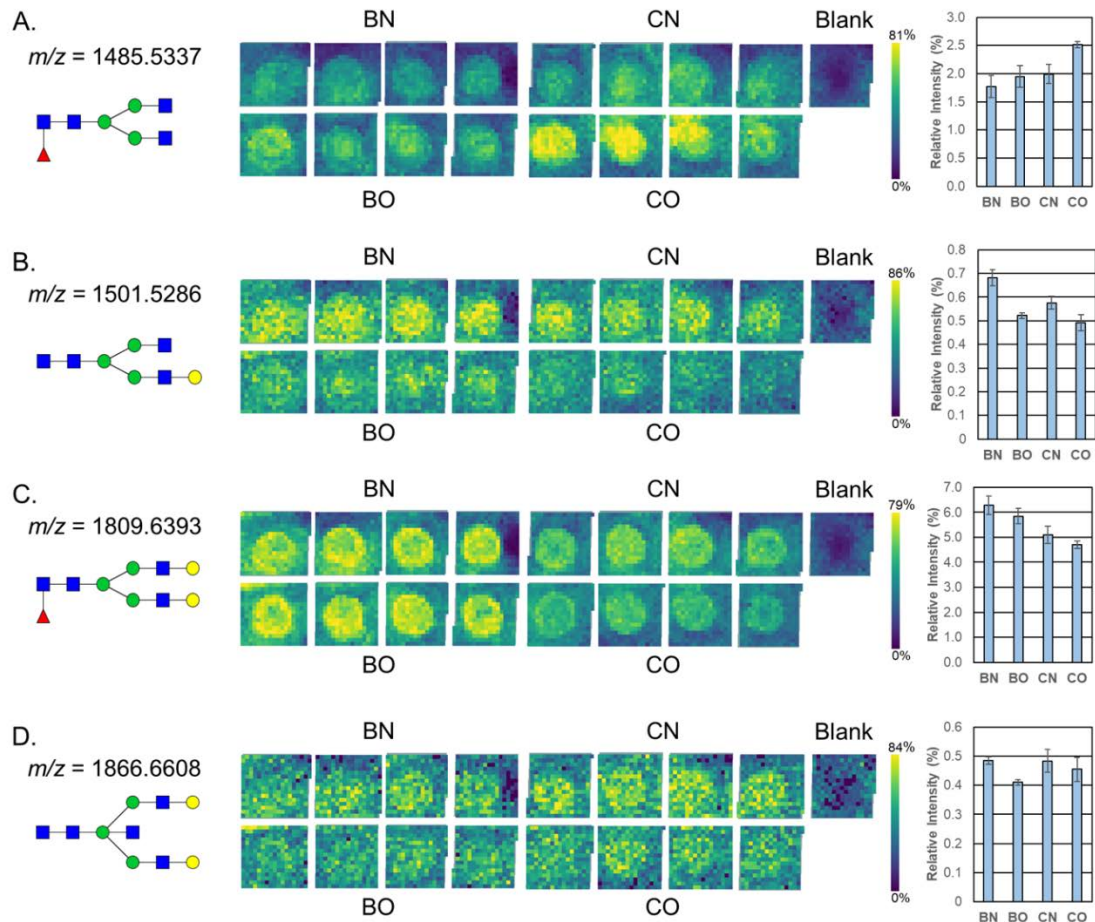


Figure 15. Application of the N-glycan profiling method to the serum of breast cancer patients. N-glycan profiles of pooled serum samples from obese and nonobese patients that had a benign lesion or breast cancer were analyzed. The average relative intensities of (A) m/z 1485.5337, (B) m/z 1501.5286, (C) m/z 1809.6393, and (D) m/z 1866.6608 N-glycans are shown. Error bars display the standard deviation for the four replicates. The relative intensities of the blank wells were used for background signal subtraction. BN = nonobese patient with a benign lesion, BO = obese patient with a benign lesion, CN = nonobese patient with breast cancer, CO = obese patient with breast cancer.

3.5 Discussion

In this study, a comparatively rapid and reproducible slide-based method for N-glycan profiling analysis of serum and plasma N-glycan profiling was presented. The approach possesses the advantages of related glycan profiling protocols for tissues, cells and antibody capture arrays that our collective group have previously reported,^{19,247,248} in that it requires minimal sample processing and no derivatization or purification of PNGase F released glycans. Using only 1-2 µl of serum or plasma, the sensitivity of this method allows for detection of a broad set of N-glycans, comparable to other N-glycan profiling methods, as well as single day analysis timelines. By repeating this workflow using the same samples on three successive days, the variation of glycan intensities day-to-day was shown to be consistent with reported technical variation of MALDI MS applied to serum.⁹¹ As was also reported,⁹¹ the use of MALDI facilitates rapid analysis of structural class groupings of N-glycans as shown in Figure 8 and Figure 9. The current study was focused more on establishing the method, however, use of a small cohort of clinical breast cancer samples was evaluated for proof-of-concept applications (Figure 15). Previous studies have examined the change of the total serum N-glycome in breast cancer patients and identified N-glycans and N-glycan structural motifs that are associated with the presence of cancer, metastasis, and cancer stage.^{206,207,216-219} Of the glycans detected in our pooled breast cancer samples, only a small set of glycans were found to have differing levels of abundance in one or more samples (Figure 15). While statistical analysis for clinical application is not possible with these samples, both

the increase of the N-glycan $m/z = 1485.5337$ and decrease of the N-glycan $m/z = 1809.6393$ have been found in previous studies.^{206,218,252} Larger cohort studies are ongoing to better identify potential serum and plasma glycan biomarker candidates related to cancer and other diseases.

The amine-reactive hydrogels that were used offered several advantages compared to other slide chemistries tested. For example, the most direct approach would be to spot serum on a glass histology slide and apply PNGase F and matrix without any other processing steps. This was done in multiple iterations, and it resulted in no glycan signal detected by MALDI-FTICR MS. Two factors are critical: First, the salts and ions present in serum and plasma lead to ion suppression. Second, the proteins are concentrated and non-denatured, limiting efficient access to PNGase F. The use of the amine-reactive gel slide provided covalent attachment of the serum and plasma proteins, which allowed denaturing and delipidating with Carnoy's solution, followed by removal of salts and ions with water washes. This covalent attachment also allows any plasma formulation to be used. Plasma with EDTA and citrate anti-coagulants have been successfully used with this workflow. Plasma with heparin anti-coagulant has not yet been evaluated, but should work like the other plasma samples, or due to its highly charged sulfation characteristics, readily adaptable to additional wash step strategies. Similarly, the workflow could serve as a starting point for analysis of N-glycans in other types of non-blood derived biofluids. This report focused specifically on use of glass slides, but efforts are ongoing to adapt the workflow to standard MALDI steel plates. It should also be noted that only CHCA matrix

was used for analyses in positive ion mode. It is certainly feasible that other matrix compounds could be evaluated, as well as use of this method in negative ion mode.

One of the goals in the method development was to simplify the sample processing steps and thereby improve the timelines of analysis. There are multiple steps in the current protocol workflow that could be further adjusted to both improve reproducibility and decrease time of analysis. The slide-based approach for sample spotting and washing could easily be adapted to sample spotting platforms and automated liquid handling systems. This could potentially decrease sample spotting and processing time, increase numbers of samples analyzed per slide, and lower CV's. While this was not reported herein, the PNGase F PRIME enzyme (N-Zyme Scientifics) is stable at temperatures up to 50 °C and results in shorter reaction times. This could decrease at least 60-90 min in the processing timeline, and is under evaluation. Additionally, the MALDI-FTICR-MS analysis reported herein used a larger sample and slide area for analysis as part of the method optimization steps. As mentioned, automated sample spotting and robotic sample handling would integrate well at this step. The mass spectrometry detection aspects of the protocol have also been assessed on the timsTOF fleX mass spectrometer (Bruker). Even using the current larger sample areas, this instrument has reduced the time for data acquisition to 30 min or less per slide. Overall, the time elapsed from sample preparation to data analysis is very quick compared to other clinical biofluid N-glycan profiling methods.⁹¹ Using the aforementioned strategies, we expect that

the current workflow can be adapted to a 4-5 hr analysis time frame from start to finish per slide.

The use of MALDI for this analysis comes with multiple limitations relative to other analytical techniques, as highlighted in the extensive multi-laboratory evaluation of HPLC, capillary electrophoresis and MALDI mass spectrometry evaluations of the same clinical serum cohort.⁹¹ In particular, MALDI MS is limited in determining the many structural isomers associated with fucosylation and sialylation, and stability of sialic acids on N-glycans is a well described issue. High vacuum MALDI and localized heat of ionization contribute to sialic acid lability for N-glycans. The MALDI-FTICR-MS used herein has a cooling gas in the source that mitigates some of this loss,^{244–246} and chemical amidation strategies have been reported that stabilize sialylated N-glycans and provide isomeric linkage information.^{91,94,253} Adapting these amidation reactions conditions to the slide-based workflow is currently ongoing. The application of a mixture of Endo F3 and PNGase F to serum and plasma, shown in Figure 12, represents an initial demonstration of feasibility for differentiating N-glycan fucose isomers within the protocol. Another modification to this workflow that can significantly increase the data acquired is utilizing an instrument like the timsTOF fleX (Bruker) that has ion mobility capabilities. Determining the collision cross section of the N-glycans is another dimension of data that could be used to confirm structure annotations and determine linkages, especially for distinguishing isomeric species. Alternatively, more broad interrogations of branched structures can be conducted using this workflow by adding sialidase to the enzyme treatment. This was shown

to increase the workflow's sensitivity to several low abundant polylactosamines, some of which were undetectable without sialidase treatment.

Chapter 4: Urine and Prostatic Fluid N-glycan Profiling by
MALDI-IMS

This chapter has been adapted from a manuscript published in *Frontiers in Chemistry*, September 2021 9:e734280. CRKB, JPH, GG, LL, and RRD carried out experiments. CRKB performed MS data processing and sample analysis assisted by RRD. OJS, JDW, JEI, CHH, and JON provided the study sample cohorts. CRKB and RRD prepared the manuscript with contributions from all authors.

**Direct N-glycosylation Profiling of Urine and Prostatic Fluid Glycoproteins
and Extracellular Vesicles**

Calvin R. K. Blaschke, Jordan P. Hartig, Grace Grimsley, Liping Liu, O. John Semmes, Jennifer D. Wu, Joseph E. Ippolito, Chanita Hughes-Halbert, Julius O. Nyalwidhe, and Richard R. Drake.

4.2 Introduction

In the search and characterization of disease biomarkers for use in liquid biopsy applications, proximal fluids like blood and urine are commonly used. Proximal fluids are found adjacent to a given tissue or organ and represent a repertoire of secreted proteins and shed cells reflective of the physiological state of that tissue. For prostate cancer and other genitourinary diseases, proximal fluids are represented by seminal plasma and expressed-prostatic secretion in urine (EPSu) ^{254,255}. EPSu, also termed post-digital rectal exam (DRE) urine, represents the fluid being secreted by the prostate following a digital rectal prostate massage, which in turn can be collected in voided urine post-exam ^{254,256}. The prostate gland secretes many proteins and other biomolecules in a prostatic fluid that combines with seminal fluid and sperm from the seminal vesicles during ejaculation. Many of these prostatic proteins are glycoproteins, like prostate specific antigen (PSA), that perform functions to activate sperm and suppress the vaginal immune micro-environment. Our group has previously characterized the proteomic composition of EPSu and prostatic secretions, identifying hundreds of different prostate-derived glycoproteins ^{116,256–258}. Development and evaluation of extensive targeted proteomic assays to these proteins in EPSu are in progress for use in prostate cancer diagnosis ^{259,260}. The prostatic fluids, EPSu and urine are also rich in extracellular vesicles (EV), which are a source for many ongoing non-coding RNA and related oligonucleotide-targeted diagnostic assays for prostate cancer and multiple diseases ^{261–263}. EV

obtained from urine and EPSu continues to be a highly active area for diagnostic assay development ^{264,265}.

The majority of proteins in EPSu, urine, and associated EVs are glycosylated ^{116,256–258}, either on asparagine residues, termed N-glycosylation, or on serine or threonine residues, termed O-glycosylation. Changes in glycosylation have been well documented in prostate cancer associated tissues, cells and biofluids ^{117,266,267}. Extensive glycoproteomic approaches, i.e., characterization of the glycan structures at the peptide sites of modification, have been reported for urine and EPSu glycoprotein targets ^{268–270}. One of the most highly characterized glycoproteins is PSA ^{125,270–274}, due to its known role in prostate cancer diagnosis and relatively simple glycosylation pattern of having a single N-glycosylation site. Although changes in glycosylation of PSA and many other prostatic glycoproteins have diagnostic potential, assays to efficiently characterize N-glycans in urine can be lengthy and require multiple processing steps, precluding large scale clinical utility ^{275,276}. Lectin arrays have been effectively used to profile glycan motifs in large cohorts of clinical urine samples ^{277,278}, however this approach cannot determine full glycan compositions or distinguish N-linked or O-linked origins. Based on an adaptation of a recently published workflow for rapid characterization of serum and plasma N-glycans ²⁷⁹, we report herein a more efficient slide-based approach combined with MALDI imaging mass spectrometry (IMS) workflows to detect total N-glycan profiles of urine, EPSu and prostatic fluid samples. Comparative results with N-glycans

detected and histologically mapped in prostate cancer tissues by MALDI IMS are also included.

4.3 Methods

4.3.1 Materials

Amicon Ultra 10k centrifugal filters were obtained from Merck Millipore (Carrigtwohill, IRL). Hydrogel coated slides (Nexterion® Slide H) were obtained from Applied Microarrays (Tempe, AZ). The rotary tool was a Dremel 200 series. The well slide module (ProPlate Multi-Array Slide System, 64-well) was obtained from Grace Bio-Laboratories (Bend, OR). Sodium bicarbonate, trifluoroacetic acid (TFA), and α -cyano-4-hydroxycinnamic acid (CHCA) were obtained from Sigma-Aldrich (St. Louis, MO). HPLC grade water, 1X phosphate buffered saline (PBS), acetonitrile, citraconic anhydride, glacial acetic acid, methanol, xylene, and chloroform were obtained from Fisher Scientific (Hampton, NH). Ethanol was obtained from Decon Labs (King of Prussia, PA). Peptide-N-glycosidase F (PNGase F) PRIME™ was from N-Zyme Scientifics (Doylestown, PA). H&E stains were obtained from Cancer Diagnostics (Durham, NC).

4.3.2 Clinical Samples

All samples were collected from patients and utilized after informed consent following Institutional Review Board-approved protocols at Urology of Virginia, Sentara Medical School, and the Eastern Virginia Medical School. All personal information or identifiers beyond diagnosis and lab results were not

available to the laboratory investigators. EPS-urine samples were collected performing a gentle massage of the prostate gland during DRE prior to biopsy, as previously described ²⁵⁶. The massage consisted of three strokes on each side of the median sulcus of the prostate and the expressed fluid from the glandular network of the prostate was subsequently voided in urine. Pools (25-50 ml/sample) of EPSu were derived from 10 patients classified as having high grade, Gleason 8-10 tumors and 10 patients with low grade, Gleason 6, organ-confined prostate cancer as described previously ¹²⁵. For isolation of EPS-derived extracellular vesicles (EPSev), the two EPSu pools (45 mL) were centrifuged at 25,000 x g for 30 minutes, and the supernatant centrifuged at 100,000 X g for 4 hours. The pelleted exosomes were washed twice with PBS and resuspended in 0.5 ml PBS, as previously described ¹²⁵. Direct EPS fluids (EPSd) were obtained under anesthesia prior to prostatectomy as previously described ²⁵⁷. A subset of 10 pairs of patient samples who provided both EPSu and EPSd were selected. Prior to glycomic analysis, 0.125 ml aliquots of each EPSu and EPSd sample were concentrated in a 10,000 MW filter cut-off 0.5 ml Amicon tube by centrifugation at 11,000 rpm in a Sorvall Legend Micro 21 benchtop microcentrifuge for 25 minutes. To each filtration tube was added 0.125 ml of 1X PBS, and centrifugation was repeated for 25 minutes. The remaining concentrated fluid, approximately 15-20 microliters, was removed to a separate vial. Each tube was rinsed with 20 microliters of PBS, and added to the concentrated sample vial (final volume 35-40 microliters). A de-identified prostate tumor tissue pair of Gleason grade 8 (4+4)/stage pT3b and patient-matched

distal non-tumor tissue was obtained from the Hollings Cancer Center Tissue and Analysis Biorepository at the Medical University of South Carolina. A serial section of each tissue was H&E-stained according to a standardized protocol.

4.3.3 Control Urine Standards

Commercial urine samples representing pooled samples from 4 healthy males and 4 females were purchased from Lee BioSolutions (Maryland Heights, MO). Prior to glycomic analysis, the control urine samples were filtered and rinsed as described for the EPS samples, except 4 ml starting volume was used with larger Amicon tubes.

4.3.4 EPS Fluids and Urine Preparation

The sample preparation and analysis of the EPS fluids (EPSu, EPSd, and EPSev) and urine samples were adapted from a workflow established for the glycomic analysis of serum and plasma²⁷⁹. After a 30-minute temperature equilibration in a moisture resistant pouch, an amine-reactive hydrogel coated slide was ground down with a rotary tool until it could fit into a Bruker MTP Slide Adapter II. A 64 well module was attached and outlined on to the back of the slide. Then the well module was unattached. Two microliters of sodium bicarbonate (100 mM, pH 8.0) was mixed with 1 microliter of the sample and briefly mixed. Within the outline of a well, 1 microliter was spotted onto the slide. EPSu, EPSd, and EPSev samples were spotted in technical triplicates, and the control urine samples were spotted in technical quadruplicates. The slide was

placed in a humidity chamber, made from a culture dish with a Wypall X 60 paper towel lining the bottom and two rolled KimWipes saturated with distilled water on opposite sides, for 1 hour on the benchtop to immobilize the samples to the slide. The slide was then dried in a desiccator for 15 minutes. The well module was reattached to the slide, matching the wells with the outlines drawn on previously. The samples were washed with Carnoy's solution (10% glacial acetic acid, 30% chloroform, and 60% 200 proof ethanol) 3 times for 3 minutes each, and subsequently washed with HPLC-grade water once for 1 minute. For the washing and rinsing steps, 50 microliters of solution was added to each well and dumped out of the well by inverting the slide. Following the water wash, the slide was dried in a desiccator for 30 minutes with the slide module attached. After detaching the slide module, a M5 TM-Sprayer (HTX Technologies) was used to spray a 0.1 mg/mL PNGase F PRIME solution in water on to the slide for 15 passes at 25 microliters/min, 1200 mm/min, 45°C, and 3 mm spacing between passes with 10 psi nitrogen gas. The slide was then incubated in a preheated humidity chamber at 37°C for 2 hours. A M5 TM-Sprayer was also used to apply the MALDI matrix solution (7mg/mL of CHCA in 50% acetonitrile/0.1% TFA) on the slide for 10 passes at 100 microliters/min, 1200 mm/min, 79°C, and 2.5 mm spacing between passes with 10 psi nitrogen gas.

4.3.5 Prostate Tissue Preparation

The tissues were prepared as described previously²⁴⁷. Briefly, the tissues were dewaxed by 1 hour in 60°C and xylene washes, rehydrated with a gradation

of ethanol and water washes, and underwent antigen retrieval in citraconic anhydride buffer (25- μ l citraconic anhydride, 2- μ l 12 M HCl, 50-ml HPLC-grade water, pH 3.0 \pm 0.5) in a decloaking chamber at 95°C for 30 minutes. A M5 TM-Sprayer (HTX Technologies) was used to spray a 0.1 mg/mL PNGase F PRIME solution in water on to the slide for 15 passes at 25 microliters/min, 1200 mm/min, 45°C, and 3 mm spacing between passes with 10 psi nitrogen gas. The slide was then incubated in a preheated humidity chamber at 37°C for 2 hours. A M5 TM-Sprayer was also used to apply the MALDI matrix solution (7mg/mL of CHCA in 50% acetonitrile/0.1% TFA) on the slide for 10 passes at 100 microliters/min, 1200 mm/min, 79°C, and 2.5 mm spacing between passes with 10 psi nitrogen gas.

4.3.6 MALDI Imaging Mass Spectrometry

A dual source timsTOF fleX MALDI-QTOF mass spectrometer (Bruker) was used to image the slides as previously described²⁸⁰. Images were collected with a SmartBeam 3D laser operating at 10,000 Hz with a 20 μ m laser spot size at a 150 μ m raster with 300 laser shots per pixel. Samples were analyzed in positive ion mode spanning a m/z range of 700-4000.

4.3.7 Data Processing and Analysis

Mass spectra were imported in to SCiLS Lab software 2021a (Bruker), normalized to total ion current, and manually peak selected for N-glycans based on theoretical mass values. SCiLS was also used for individual peak visualization

and quantification. Maximum mean values for each peak were exported for each sample region. Each N-glycan measurement for each sample was subtracted by the background signal in the blank well to find the absolute intensity. To account for differences in protein concentrations that could lead to higher signal intensities and detection of more low-abundance N-glycan species, N-glycan relative intensities were calculated as the absolute intensity divided by the sum of all the absolute intensities of the N-glycans found in each of the samples being compared. Comparisons of the number of N-glycans detected in each sample is also discussed, and the presence/absence of a N-glycan in each sample is noted in Supplemental Table 1. N-glycan structures were labelled with a N-glycan class or classes depending on their putative structures. Quantifications of the N-glycan classes were calculated by summing the relative intensities of the individual N-glycans belonging to each class. N-glycan profiles were also examined by grouping each N-glycan into a group depending on the presence and/or absence of mannose, fucose, sialic acid, and sulphate and comparing the summed relative intensities of the classes. When comparing individual N-glycan intensities across samples, the multiply sodiated species of sialylated and sulfated N-glycans were added together.

4.4 Results

To complement previous proteomic studies of proximal prostatic fluids obtained in the urology clinic as related to prostate cancers^{116,256–258} a series of different EPSu, EPSd and EPSev samples were used to develop a MALDI-based N-glycan profiling method. The goal was to have a workflow that required minimal sample processing and could be completed in a 6-8 hour timeline, in contrast to current glycomic analysis workflows for urine and prostatic fluids that require multiple processing, derivatization and purification steps prior to analysis. A previous slide-based approach used for serum and plasma N-glycan profiling²⁷⁹ was the starting point, and a workflow summarized in Figure 16 was developed for EPSu. A key feature is use of an amine reactive slide chemistry that covalently binds target glycoproteins, facilitating washing steps to remove lipid and salts prior to spraying of a molecular coating of PNGase F PRIME to release N-glycans. An additional concentration and buffer exchange step was added for EPSu and EPSd, using a 10000 MW cut-off spin cartridge to concentrate and allow buffer exchange of the sample prior to addition to the amine-reactive slide. This step results in a 3-4 fold increase in concentration of glycoproteins in the biofluid. An SDS-polyacrylamide gel image showing protein loading examples for the EPSu, EPSd and EPSev samples are provided in Figure 17.

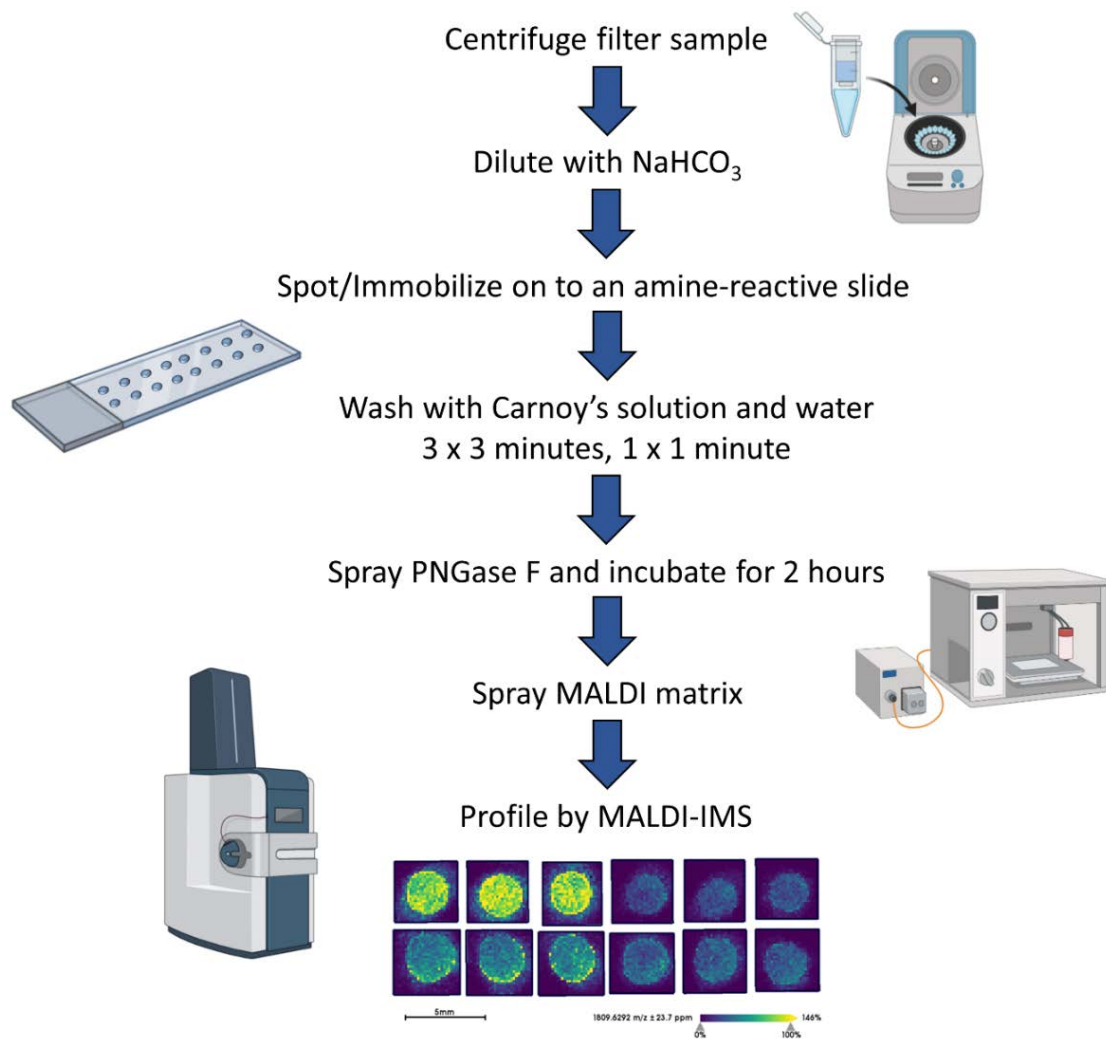


Figure 16. Workflow for slide-based urine and prostatic fluid N-glycan profiling by MALDI IMS.

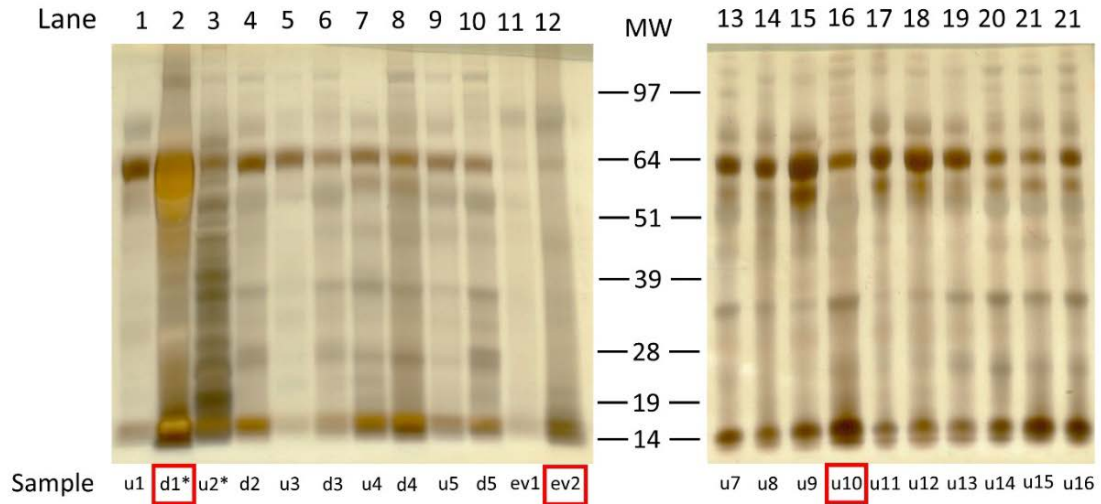


Figure 17. SDS gel separation of the samples. One microliter (approximately 10-15 ug protein) of each sample was added in 10 ul SDS loading buffer and separated on a 10% Genescript ExpressPlus polyacrylamide gel in MOPS buffer. SeeBlue Plus2 pre-stained molecular weight standards from Invitrogen were also loaded. Proteins were visualized by silver staining. EPSu (labelled with a “u”) and EPSd (labelled with a “d”) sample pairs, and EPSev (labelled with “ev”). Sample labels with an asterisk were used for mass spectra comparison, and sample labels with a red box were used as representative samples.

Initial experiments focused on optimizing detection of N-glycans in EPSu and EPSd sample pairs obtained from the same donors, as well as existing EPSev pool samples. Similar to what was previously determined for serum or plasma preparation on the amine-reactive hydrogel slides, the key for optimal N-glycan detection was inclusion of the Carnoy’s solution wash after spotting, which

serves to remove lipids and salts, as well as denature the bound glycoproteins facilitating access for PNGase F. Thus far, a total of 35 EPSu, 10 EPSd and 8 EPSev samples have been analyzed with the workflow shown in Figure 16. Cumulatively, the resulting N-glycans detected in each sample type are summarized in Table 2, and structural class groupings are shown in Figure 18. Broadly, the EPSu samples had the most N-glycan species detected (n = 182) versus EPSd (n=135). These numbers include multiple versions of the same N-glycan compositions for sialylated and sulfates species, which can vary in mass due to varying numbers of sodium ions associating with the charged groups. These glycoforms were included in Table 2 but were generally detected at lower intensity values.

There was a range of protein concentrations across the samples in each sample type examined. While this created differences in the total intensity of the N-glycan profile and number of N-glycans that could be detected, these differences were accounted for by only comparing relative intensities, i.e. an individual N-glycan's intensity relative to the total intensity of the N-glycans in that sample that were also seen in all sample types being compared. Representative samples with the most N-glycans detected were selected and compared from each sample type to display the breadth the N-glycan profiles.

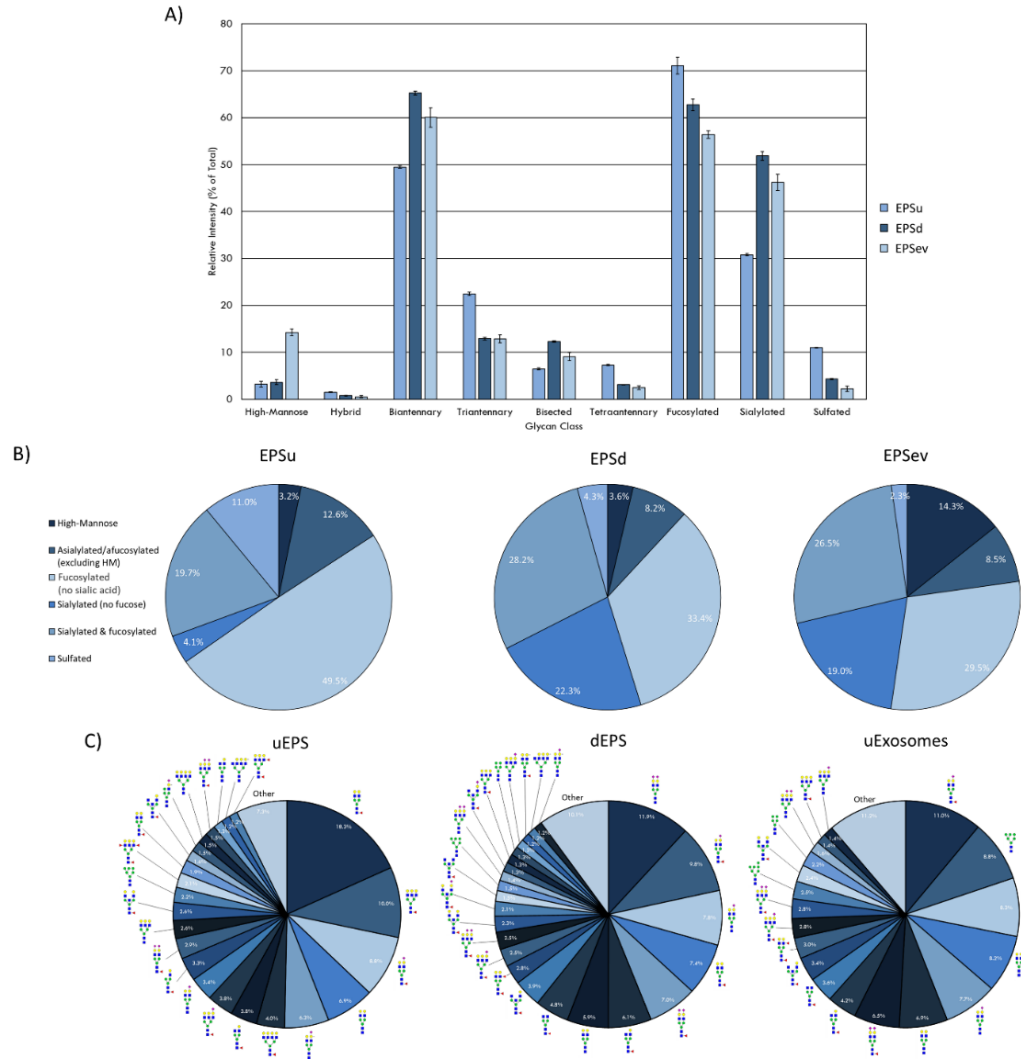


Figure 18. N-glycan profile of prostatic fluids (EPSu, EPSd, and EPSev). (A) Glycan class abundances of representative samples determined by summing the relative intensities of the N-glycans assigned to that class. Error bars represent the standard deviation. (B) Distributions of representative samples' N-glycans assigned into groups based on composition. (C) Distribution of total N-glycan intensity by individual N-glycan, with N-glycans comprising at least 1% of the total intensity displayed and annotated.

The intensity of the N-glycan classes varied across the prostatic fluid samples (Figure 18 A). EPSd and EPSev N-glycan classes had similar intensities, except for the higher amount of high-mannose N-glycans in EPSev. For all samples, the majority of the N-glycans were biantennary and/or fucosylated. EPSu had approximately 15% and 20% less sialylation than the EPSd and EPSev, respectively, but had an increased amount of tetraantennary and sulfated N-glycans. Many of these findings were replicated when grouping the N-glycans detected in each sample based on composition (Figure 18 B). About half of the N-glycans in EPSu were fucosylated with no sialic acid compared to approximately 30% in the EPSd and EPSev samples. The EPSd and EPSev samples had higher levels of N-glycans with sialic acids and no fucose. When examining the intensity of individual N-glycans in the prostatic fluid samples, the most abundant N-glycans were typically biantennary with two galactoses (Figure 18 C). In concordance with the N-glycan class comparison, EPSev had more high abundance high-mannose N-glycans than the other samples, and m/z 1419.4755 (Hex6HexNAc2 + 1Na) was the second most abundant N-glycan. The sulfated N-glycan m/z 2056.6156 (Hex5HexNAc4NeuAc1 + 1SO4 + 2Na) had a relative intensity of 6.3% in the EPSu, compared to 1.2% in EPSd and less than 1% in EPSev.

In order to compare mass spectra and absolute intensity values of N-glycan peaks, an EPSu and EPSd sample with similar protein levels from the same cohort were prepped together and imaged in the same run (Figure 19). Substantial differences were seen in most of the high-intensity N-glycan peaks,

except for m/z 1809.6393 (Hex5dHex1HexNAc4 + 1Na), m/z 2465.8669 (Hex6dHex1HexNAc5NeuAc1 + 1Na), and m/z 2800.9263 (Hex6dHex1HexNAc5NeuAc2 + 3Na). The increased levels of sialylated N-glycans in the EPSd sample and increased fucosylated species in the EPSu sample is also evident.

The advantage of developing the assay using EPSu samples are the inherently higher protein concentrations present from the prostatic fluid mixture in these samples relative to normal urine. Therefore, for comparison, control urine samples from pools of 4 healthy male and 4 healthy female donors were processed and analyzed using this workflow. This required 4 ml of starting fluid for concentration and desalting using larger filtration tubes, but was otherwise the same workflow as described for EPSu/d samples. Overall, N-glycan class intensities were similar for both samples (Figure 20 A). There were high levels of biantennary, fucosylated, and sialylated N-glycans and low levels of hybrid N-glycans. The rest of the N-glycan classes had approximately 10% intensity. The biggest difference between the male and female samples was the slightly lower level of fucosylation in the female sample. The distribution of N-glycans based on composition displayed very similar profiles (Figure 20 B). The biggest difference was a 3.2% increase in high-mannose N-glycans in the female sample. Accordingly, m/z 1419.4755 (Hex6HexNAc2 + 1Na) was the sixth most abundant N-glycan in males and the second most abundant in females (Figure 20 C). Similar to the EPSu samples, the most abundant N-glycans are primarily biantennary with two galactoses.

As the glycoprotein constituents of the EPSu and EPSd samples are secreted by prostate glands, a comparative N-glycan comparison of prostate tissues was done using previously reported N-glycan MALDI IMS approaches^{245,280,281}. In the example shown in Figure 21 A, a pair of non-tumor and tumor tissues from the same donor were evaluated. Use of the non-tumor tissue from the same donor is done to demonstrate the specificity of tumor specific N-glycans detected in the tumor containing tissue (a Gleason grade 8 (4+4), stage pT3b). The tumor region is quite distinct, localized to the bottom left corner of the tissue. Additional H&E images of both tissues with increasing magnification are provided in Figure 22. In Figure 21 B and C are shown representative N-glycan classes and their tissue distributions, from a total of 73 N-glycans detected. These structures include tumor-associated paucimannose (Hex3HecNAc2), high mannose (Hex5-Hex9HexNAc2), and branched fucosylated species (Hex7HexNAc6Fuc2)(Figure 21 C), consistent with previous reports²⁸¹. Two of the most abundant sialylated biantennary N-glycans detected are stroma-associated, as shown in the overlay image with a tumor N-glycan (in red), Hex5HexNAc4NeuAc1 in blue and Hex5HexNAc4Fuc1NeuAc1 in green (Figure 21 B, and individually in Figure 21 C). A segmentation analysis of the 73 N-glycans is shown in Figure 21 D, illustrating how different N-glycan classes are associated with different histopathology features.

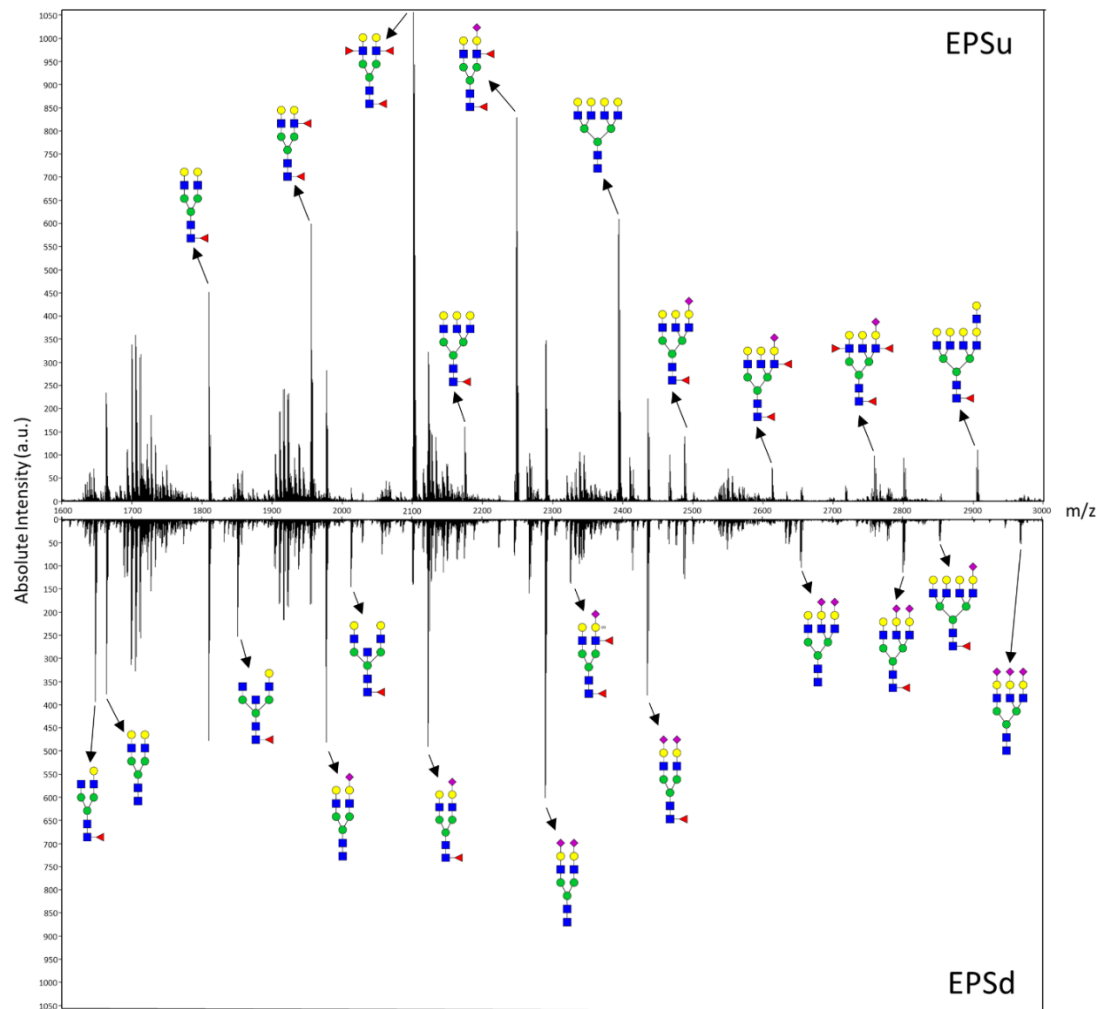


Figure 19. Mass spectra of representative EPSu (top) and EPSd (bottom) samples with annotations for a set of putative N-glycan structures. The N-glycan compositions are represented by blue squares for N-acetylglucosamine, green circles for mannose, yellow circles for galactose, purple diamonds for sialic acid, red triangles for fucose, and an “S” for sulfate.

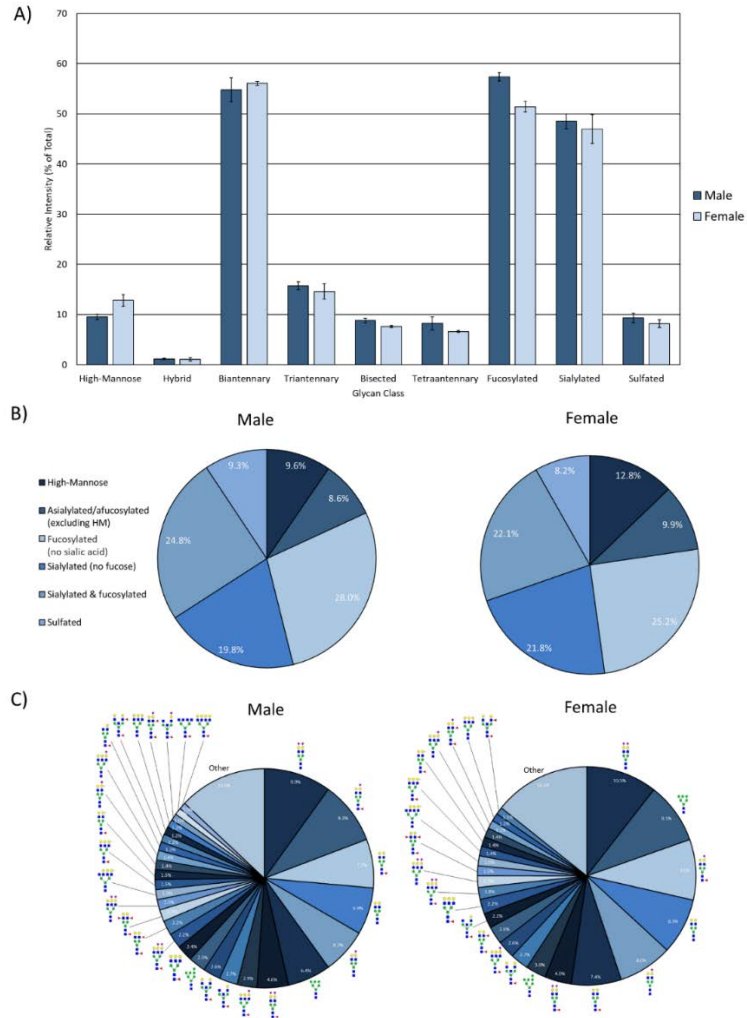


Figure 20. N-glycan profile of pooled healthy male and female urine analyzed by MALDI IMS. (A) Glycan class abundances of representative samples determined by summing the relative intensities of the N-glycans assigned to that class. Samples were analyzed in technical quadruplicate. Error bars represent the standard deviation. (B) Distributions of representative samples' N-glycans assigned into groups based on composition. (C) Distribution of total glycan intensity by individual N-glycan, with N-glycans comprising at least 1% of the total intensity displayed and annotated.

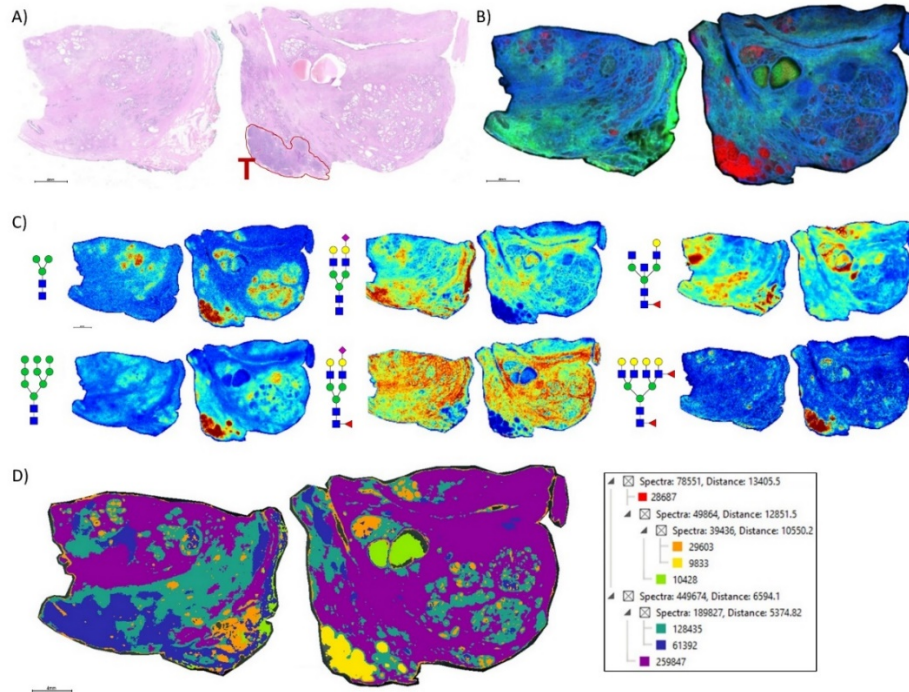


Figure 21. N-glycan imaging of FFPE prostate cancer and non-cancer tissues.

(A) H&E stain of the two tissues from the same donor, non-cancer tissue on the left side and tumor on the right side. The tumor region, a Gleason grade 8/stage p3Tb, is highlighted with red outline and a red T. (B) An overlay MALDI image of three N-glycans, two stroma-associated ones, Hex5HexNAc4NeuAc1 in blue and Hex5HexNAc4Fuc1NeuAc1, and a tumor glycan Hex7HexNAc6Fuc2 (in red). (C) Six representative individual N-glycan images representative of different structural classes are shown, and glycan structure: Hex3HecNAc2, Hex5-Hex9HexNAc2, Hex5HexNAc4NeuAc1, Hex5HexNAc4Fuc1NeuAc1, Hex4HexNAc4Fuc1 and Hex7HexNAc6Fuc2. (D) Segmentation analysis, using Manhattan and k-bisecting classifications, was applied to the 73 N-glycans detected in the tissues. Spectra groupings are shown in the adjacent data tree.

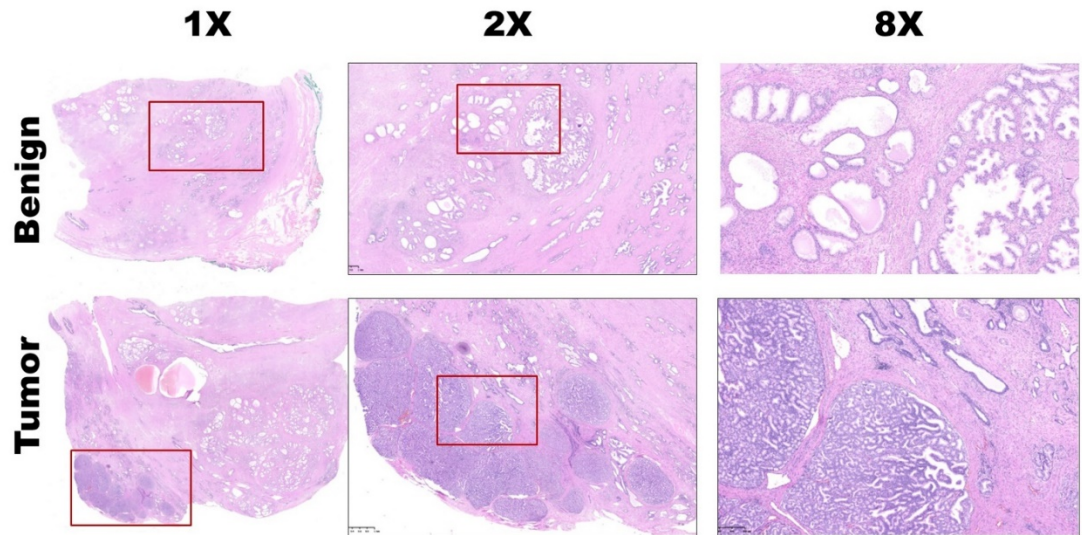


Figure 22. Higher resolution images of the H&E-stained prostate tissues. Images were obtained following H&E staining of each slide and scanning into a Hamamatsu Nanozoomer 2.0RS digital slide scanner. Images of 2X and 8X are shown for selected regions, highlighted in the preceding image with a red box outline.

This representative tissue was selected for another feature, as it was noted that there was a distinct intra-lumen glandular N-glycan signature that could be detected. As shown in Figure 23 A, the two sialylated biantennary N-glycans provide a stromal scaffold image. This was used to detect which N-glycan species were present in the lumen regions, as illustrated in Figure 23 B, and in Figure 23 C for a highlighted gland region. Doing this, 38 N-glycans were detected in the lumen of glands.

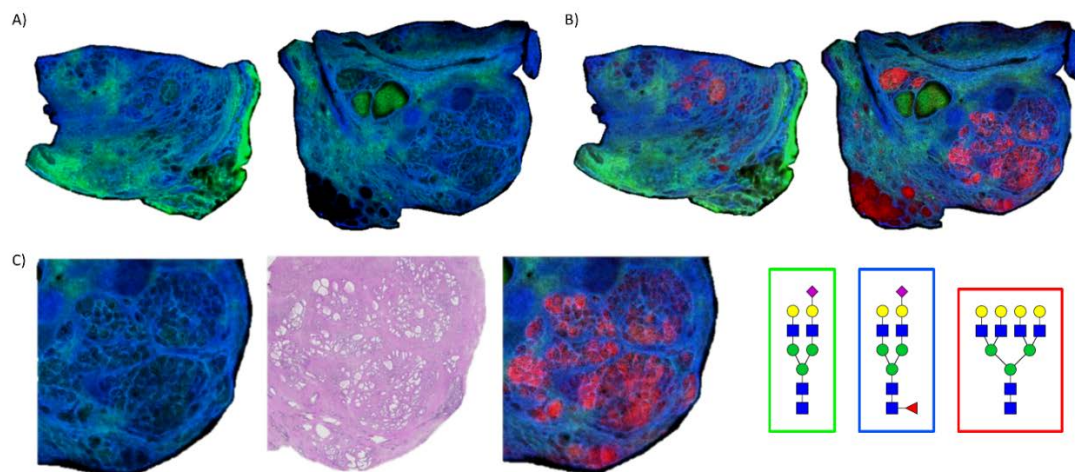


Figure 23. Detection of intraluminal N-glycans in non-tumor prostate glands. (A) Two distinct stroma-associated glycans, Hex5HexNAc4NeuAc1 in blue and Hex5HexNAc4Fuc1NeuAc1 in green, were chosen to highlight the locations of the prostate glands. (B) An example of a luminal glycan is shown in red (Hex7HexNAc6), overlaid with the stroma glycans. (C) Highlighted region of the tumor tissue, bottom right corner, to illustrate the glycan distributions inside the glands. The analogous region in the H&E stained slide is also shown. Structures are bordered with the color that matches their tissue localizations.

Using the lists of 73 tissue N-glycans and the subset of 38 lumen N-glycans, these were compared to the N-glycans detected in EPSu and EPSd samples. As shown in the Venn diagram in Figure 24, 44 N-glycans were present in the tissues and both EPS samples (Figure 24 A), and 18 N-glycans were shared in the lumen and EPS samples (Figure 24 B). Structurally, these shared N-glycans were the high mannose and the most abundant branched fucosylated N-glycans.

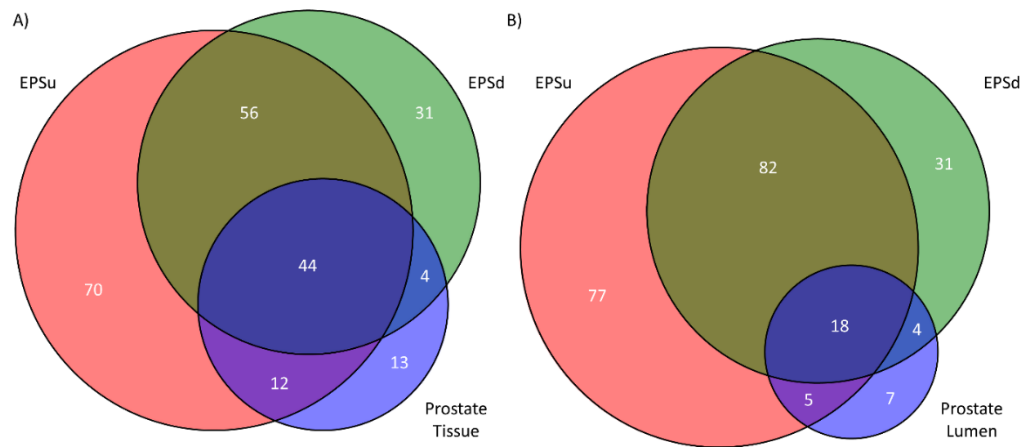


Figure 24. Comparison of detected N-glycans across sample sets. (A) The number of unique and common N-glycans detected in EPSu, EPSd, and across the entire prostate tissues. (B) The number of unique and common N-glycans detected in EPSu, EPSd, and within the lumen of the prostate tissues.

4.5 Discussion

The goal of our study was to develop and optimize a more efficient and rapid method to evaluate N-glycosylation by MALDI-MS in urine and EPSu samples. The workflow is similar to what was developed for serum and plasma ²⁷⁹, utilizing an amine reactive slide to capture glycoproteins on a solid surface. An additional 1 hour step of sample processing by molecular weight filtration and centrifugation for urine was included in order to concentrate and solvent exchange. Spotting only 1 ul, the urine glycoprotein spots are sprayed with a molecular coating of PNGase F, followed by analysis on a MALDI-FTICR or MALDI-QTOF mass spectrometer. The method is applicable to urine, EPSu and EPSd samples. An EPSev prep was spotted directly onto the amine reactive slides without any pre-processing. Routinely, 100 or more N-glycans can be detected depending on the sample type.

The N-glycan profiling of the prostatic fluid samples displayed more high-mannose N-glycans in the EPSev compared to the EPSu and EPSd, reflected in the intensities of the N-glycan classes and the most abundant individual N-glycans (Figure 18). Relatively high intensities of high-mannose N-glycans have been noted in non-EPS urinary exosomes as well ²⁸². The similarity of the EPSd and EPSev N-glycan profiles suggests that a significant portion of the EPS glycoproteins may be within exosomes. In conjunction with these findings, direct comparison of mass spectra from EPSu and EPSd with similar protein levels displayed distinct N-glycan profiles for EPSu and EPSd (Figure 19). While male and female urine N-glycan profiles and individual N-glycan abundances

displayed few differences, there was a slightly higher amount of fucosylation in the male sample. This trend has previously been identified in male and female plasma samples ⁷⁷. Similar levels of sulfation were detected in the urine controls and EPSu, which was higher than the EPSd and EPSev, indicating that the sulfated N-glycans in EPSu may be originating from urine glycoproteins.

A comprehensive MS analysis of the N-glycans present in adult and pediatric urine samples has been published ¹¹². In this study, PNGase F released N-glycans were isolated, labeled with aminobenzoic acid, and methylamidated prior to analysis by LC-MS/MS. The authors reported 116 N-glycan compositions could be detected, and a subset of 46 N-glycans that were reproducibly detected and further quantified ¹¹². This study provides an optimal benchmark for our current study to evaluate how many of the 46 N-glycans could be detected by our solid-phase and single day MALDI MS workflow. Comparing the MALDI data from the male and female urine standards, the majority of N-glycans are detected, especially the biantennary, high mannose and fucosylated species. The MALDI-based assay is not as effective for detection of larger tri- and tetra-sialylated N-glycans, and for N-glycans with molecular masses above 3700 m/z.

One observation noted from our data was that the intensity levels of N-glycans detected were variable, even when protein concentrations were comparable. There was also not a large variation in the types of N-glycan structures detected in the samples, especially for the most abundant species. These variations could reflect significantly different protein levels of individual glycoprotein species, or variations in the amount of the most abundant protein in

human urine, uromodulin. Uromodulin, also called Tamm–Horsfall protein, has eight N-glycosylation sites and has been comprehensively studied for N-glycosylation content and composition ¹¹³. Interestingly in this study, N-glycan analysis comparisons of a uromodulin depleted urine versus non-depleted sample indicated little difference in the N-glycan compositions detected in both samples ¹¹³. Our study only represents a few sample numbers, so no conclusions can be reached until statistically relevant sample numbers are analyzed. The rapidity and efficiency of the slide based MALDI assay will facilitate evaluation of larger cohorts. Possibly pairing the N-glycan data with the quantitative MRM proteomic assay developed for EPSu proteins could address both the changes in N-glycan levels and protein concentrations in the same sample ^{259,260}. It also does not preclude direct targeting of specific glycoproteins present for N-glycan content analysis, as is done so often for PSA. We have already reported a solid-phase antibody array approach to N-glycan profile individual serum glycoproteins captured by their specific antibodies ^{20,248}. A similar approach targeting the abundant urine and EV glycoproteins like PSA is in progress.

The N-glycans obtained from a representative FFPE prostate cancer tissue were also included to illustrate the N-glycans that can be detected in the lumen of the prostate glands, representing prostatic fluid still present during tissue fixation. Based on hundreds of prostate cancer tissues analyzed by N-glycan IMS, to be reported separately, detection of luminal N-glycans in these tissues is highly variable. Fluid remnants can be seen in the H&E stains, but not

all gland lumens contain this, and can be completely absent in many tissues. Presumably this variation reflects how much fluid was present at the time of prostatectomy but could also reflect differences in FFPE tissue preparation and processing. While this is a variable that may lack clinical diagnostic significance, the selected tissue in Figure 23 highlights the tumor-associated and secreted N-glycans (n=38) in tissue that can potentially be detected in the EPSu, EPSd and EPSev samples. These are primarily paucimannose, high mannose and branched multi-fucosylated N-glycans.

Chapter 5: Classifying Suspicious Mammogram Findings by
Serum and IgG N-glycan Profiling

This chapter was adapted from a manuscript submitted to *Scientific Reports*, September 2022. CRKB and RRD conceived the experiments. CL supervised clinical sample collection and clinical data. CRKB conducted the experiments with consultation with RRD, PMA, and ASM. CRKB and EGH analyzed the results.

Integrating age, BMI, and serum N-glycans detected by MALDI imaging mass spectrometry to classify suspicious mammogram findings as benign lesions or breast cancer

Calvin R. K. Blaschke, Elizabeth G. Hill, Anand S. Mehta, Peggi M. Angel, Christine Laronga, and Richard R. Drake.

5.2 Introduction

Breast cancer has the highest incidence rate of all cancers, and is the second leading cause of cancer deaths in US women.¹⁴⁸ The stage of the breast cancer at diagnosis has a large impact on the prognosis of the patient, highlighting the need for early detection strategies.¹⁴⁸ Accordingly, the implementation of mammography as a screening tool has decreased breast cancer mortality.²³⁰ While mammograms have a high cancer detection rate, issues of overdiagnosis and false-positives continue to affect their clinical utility.²⁸³ Due to the similar presentations of benign lesions and breast cancer in a mammogram, abnormal interpretations are followed up with additional imaging or a biopsy. A study of over 1.6 million mammograms found that 1.2% of mammograms were classified as suspicious or highly suggestive of malignancy, but after the recommended biopsy, only 71.6% of those patients had breast cancer.¹⁷¹ With over 39 million mammograms performed in the US every year, this results in over 345,000 unnecessary biopsies yearly in the US alone.¹⁷² To improve breast cancer screening, a supplemental assay is needed to aid in the discrimination of benign lesions and breast cancer.

While mammography is an essential tool in breast cancer screening, issues of low specificity frequently result in unnecessary biopsies or follow-up imaging. These false positives have significant financial costs on a national and individual patient level, as well as a psychological cost for the patient.^{174,284–286} A minimally or non-invasive biomarker for breast cancer detection is needed to increase the specificity of breast cancer screening. Alternative imaging modalities

have displayed improvements in cancer detection, especially for women with a higher breast density, but suffer from increased radiation exposure, high costs and low availability, or similar issues with false-positives for digital breast tomosynthesis (DBT), magnetic resonance imaging (MRI), and ultrasound (US), respectively.^{175,287–292} Using blood-based biomarkers for screening benefits from cost-effective and minimally invasive collection. The currently used blood-based biomarkers for breast cancer, such as cancer antigen (CA) 15-3, CA 27.29, carcinoembryonic antigen (CEA), and soluble human epidermal growth factor receptor 2 (sHER2), are primarily used for prognosis, staging, and therapy monitoring, but have limited utility in detection.^{178,179,189–193,180–187} These limitations prompt an investigation into other blood-based biomarkers that can detect differences in the systemic response to breast cancer and benign lesions.

Serum is a rich source of potential biomarkers because of the variety of molecules circulating in the body with altered abundance, activation, and/or composition due to the effects of a disease. Serum biomarkers can be readily adapted into clinical assays because non-invasive collection is cost-effective and large cohorts are available for thorough and accurate analysis of diagnostic potential. The composition of serum collected during blood draws reflects a dynamic biofluid comprised of thousands of proteins, lipids, metabolites, and nucleic acids. A large portion of serum consists of glycosylated proteins produced by the liver and immunoglobulins secreted by B-cells.⁷² Glycosylation is a highly regulated metabolic process shown to have essential roles in protein folding, molecular trafficking, protein clearance, and many other

processes.^{242,243,293–295} Being closely aligned with function, altered glycoforms of serum proteins can be dynamic indicators of systemic responses to disease and have been interrogated for potential clinical applications for decades.^{75–77}

The creation and optimization of high-throughput techniques for the profiling of serum N-glycans, i.e. a class of glycans linked to a protein via an asparagine residue, has enabled the interrogation of large sample cohorts for robust detection of N-glycan alterations associated with diseases and conditions.^{91,93} Recently, our group has adapted matrix-assisted laser desorption/ionization (MALDI) imaging mass spectrometry techniques for analysis of serum, urine, prostatic fluids, cultured cells, and antibody-captured serum proteins.^{18–20,279,296,297} Here, we utilized a high-throughput MALDI imaging mass spectrometry technique²⁷⁹ to identify significant differences in the serum N-glycan profiles of 199 women with benign lesions from 99 women with breast cancer. The relative intensity of individual N-glycans and N-glycan classes were compared and evaluated for the ability to discriminate between benign and cancerous conditions, with and without stratification by age and BMI.

5.3 Methods

5.3.1 Serum Samples

Donors were women who reported to the breast surgical clinic at the Moffitt Cancer Center & Research Institute for a breast biopsy after a Breast Imaging Reporting and Data System (BI-RADS) 4 imaging designation. The samples were collected after H. Lee Moffitt Cancer Center & Research Institute

institutional review board approval and informed consent from the donors. All experiments were performed in accordance with this protocol following the approved, relevant guidelines and regulations. Serum samples were collected immediately prior to tissue biopsy and deidentified. Age, body mass index (BMI), ethnicity/race, and pathology results of benign lesions (fibroadenomas, fibrocystic, etc.) and breast cancers (invasive or ductal carcinoma in situ (DCIS)) were linked with each specimen. Patients with both an invasive breast cancer and DCIS were treated as invasive. Patients self-identified their ethnicity, and patients identifying as non-Hispanic were asked to self-identify race. The benign samples (n=199) and breast cancer samples (n=99) were stratified by age groups according to the current United States Preventative Services Task Force mammogram recommendations (younger than 40, 40 - 49, 50 - 74, older than 74) and the established BMI classifications (less than 18.5, 18.5 - 24.9, 25 to 29.9, more than 29.9).²⁹⁸ The clinical characteristics of the samples analyzed in this study are summarized in Table 3.

		Benign (n=199)	Cancer (n=99)
Ethnicity / Race	Caucasian	151	85
	Hispanic	38	10
	African American	8	1
	Asian	2	3
BMI	Mean \pm SD	28.3 \pm 5.9	28.7 \pm 6.19
	< 18.5	1	2
	18.5 - 24.9	68	21
	25 - 29.9	60	42
	> 30	70	34
Age	Mean \pm SD	56.7 \pm 11.3	57.3 \pm 11.9
	< 40	14	7
	40 - 49	35	17
	50 - 74	140	68
	> 74	10	7
Pathology	DCIS	-	30
	Invasive	-	69

Table 3. Clinical characteristics of patients by diagnosis. BMI = Body Mass Index. SD = Standard Deviation. DCIS = Ductal Carcinoma In Situ.

5.3.2 Materials

Trifluoroacetic acid (TFA), sodium bicarbonate, and α -cyano-4-hydroxycinnamic acid (CHCA), octyl- β -D-glycopyranoside (OGP), bovine serum albumin (BSA) were purchased from Sigma-Aldrich (St. Louis, MO). Ethanol was purchased from Decon Labs (King of Prussia, PA). Acetonitrile, phosphate buffered saline, HPLC grade water, citraconic anhydride, xylene, methanol,

glacial acetic acid, and chloroform were purchased from Fisher Scientific (Hampton, NH). The amine-reactive slides (Nexterion® Slide H) were purchased from Applied Microarray (Tempe, AZ). The attachable well chambers (ProPlate Multi-Array Slide System, 64-well) were purchased from Grace Bio-Laboratories (Bend, OR). Peptide-N-glycosidase F (PNGase F Prime) was purchased from N-Zyme Scientifics (Doylestown, PA). Antihuman IgG was obtained from Bethyl Laboratories (Montgomery, TX).

5.3.3 Serum Sample Preparation

Serum sample preparation was performed as previously described.²⁷⁹ Serum samples were diluted in sodium bicarbonate and spotted onto an amine-reactive slide. Each slide had spot of PBS diluted in sodium bicarbonate added as a blank. Additionally, a standard healthy serum sample was added to each slide for normalization across slides. All samples and standards were spotted in technical triplicate. The slide was then placed into a humidity chamber for 1 hour to bind the serum proteins to the slide. Well chambers were attached to the slide to isolate each sample, and the samples were washed with Carnoy's solution and water to remove lipids and salts, respectively. After drying the slides, PNGase F was sprayed across the slide with an automated sprayer (M5 TM-Sprayer, HTX Technologies, Chapel Hill, NC) and incubated in a humidity chamber at 37°C for 2 hours to enzymatically cleave the N-glycans from the captured serum glycoproteins. The HTX M5 sprayer was then used to spray CHCA matrix across the slide.

5.3.4 IgG Sample Preparation

The IgG N-glycan analysis of the samples in this cohort was based on the protocol described previously.²⁰ Antihuman IgG antibody was spotted onto an amine-reactive slide in 0.75 uL spots of 150 ng/nL antibody in PBS. The antibody spots were immobilized to the slide after a 1-hour incubation in a humidity chamber. The slide was dried in a desiccator for 15 minutes, a 24-well chamber module was attached to the slide, and the slide was washed with a 0.1% OGP in PBS solution for 1 minute. To block the rest of the slide, a 2% BSA/0.1% OGP in PBS solution was added to each well, and the slide was gently shaken for 1 hour. After the blocking, the wells were washed with two 2-minute washes of PBS and a 1-minute water wash. After desiccating the slide for 30 minutes, 1 uL of serum was diluted in 99 uL of PBS, and the diluted samples were added to the wells. For each slide, 3 antibody spots had a blank of only PBS added to the well. The slide was placed in a humidity chamber with gentle shaking for 2 hours to allow for antibody capture of the IgG in the serum samples. The wells were washed twice with 0.1% OGP in PBS for 1 minute, twice with PBS for 3 minutes, and twice with water for 1 minute. After the final washes, the slides were desiccated until dry, detached from the well modules, and sprayed with PNGase F and CHCA using the HTX M5 sprayer in the same fashion as previously described for the serum sample preparation. An overview of this workflow is shown in Figure 25.

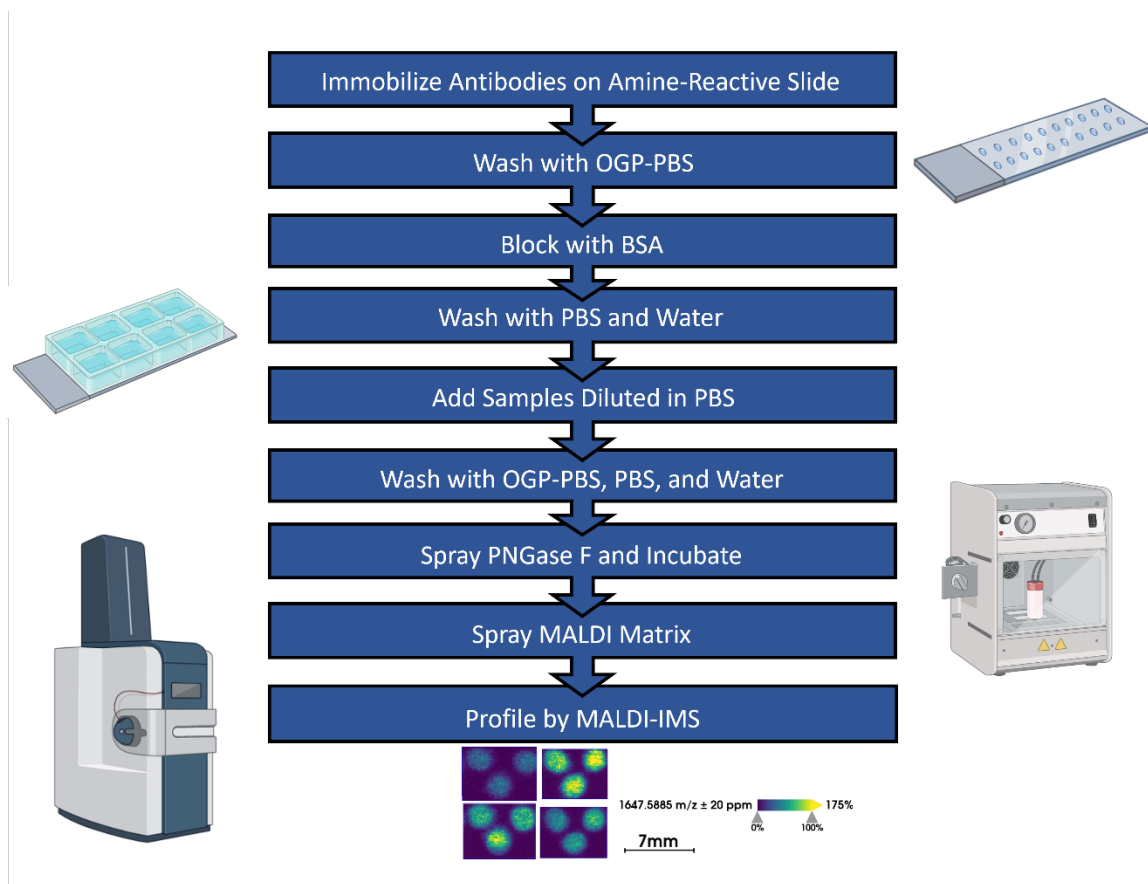


Figure 25. IgG analysis workflow overview.

5.3.5 Mass Spectrometry Data Collection

The slides were imaged with a timsTOF fleX MALDI-QTOF mass spectrometer (Bruker, Billerica, MA) with a SmartBeam 3D laser operating at 10,000 Hz and 20 μm laser spot size. There were 300 laser shots collected per pixel, and a 150 μm raster was used. A 700-4000 m/z range was scanned in positive ion mode.

5.3.6 Data Processing

The mass spectra were imported into SCiLS Lab software (2022b Pro, Bruker, Billerica, MA). Total ion current was used to normalize the images. N-glycan peaks were selected manually using theoretical mass values. Putative structures were assigned to N-glycan peaks based on previously reported assignments.^{98,280,299} N-glycans are reported as m/z values and using the Oxford nomenclature for the putative structures, where A represents the number of antennae present, F indicates the fucose, B indicates the presence of a bisecting N-acetylglucosamine, G represents galactoses, and S denotes sialic acids.³⁰⁰ N-glycan compositions and mass error are reported in Table 4. Maximum peak values were extracted and used for N-glycan absolute intensities. For each slide, the intensity of the blank sample (only sodium bicarbonate solution for serum sample analysis, the antibody spots with no serum added for the IgG sample analysis) was subtracted from the N-glycan intensity across the samples. N-glycans with less than 2 arbitrary units (au) of intensity after blank subtraction were converted to 1 au to avoid 0 values for the statistical analysis. If a N-glycan was less than 2 au for more than 20% of samples, it was excluded from further analysis. Next, the standard sample was used to create normalization factors for each N-glycan on each slide, where the intensity of the individual slide's standard was divided by the average intensity of the standards across all slides. Each slide's N-glycans were then multiplied by the corresponding normalization factor. The intensities were then converted to relative intensities by dividing individual N-glycan intensities by the summed intensities of all the N-glycans in that sample.

N-glycan class intensities were calculated by summing the relative intensities of the N-glycans containing the structural or compositional traits of the N-glycan class.

5.3.7 Statistics

Linear regression was used to model log-transformed N-glycan relative intensities as a function of disease status, controlling for age, ethnicity/race, and BMI. Logarithmic transformation was used to induce approximate normality and stabilize variance. Differences in log-relative intensities across disease status levels were evaluated using model-based contrasts. A 5% false-discovery rate was used to adjust for test multiplicity.^{301,302} A similar modeling approach was used to evaluate the association between N-glycan intensities and age or BMI. Loess smoothing was used for graphical exploration of the functional relationship prior to modeling.³⁰³ The performance of N-glycans to classify disease status was evaluated using the area under the receiver operating characteristic curve (AUROC) of a logistic regression model of the patient pathology as a function of the selected N-glycans, N-glycan classes, and clinical characteristics. AUROCs are reported with 95% confidence intervals. Optimal sensitivity and specificity were defined as the sensitivity and specificity of the threshold at the cutoff point of the curve closest to (0,1). All statistical analyses were performed using R version 4.1.0.

5.4 Results

Serum N-glycan profiles of women with either a benign lesion or breast cancer detected after a mammogram were determined by MALDI imaging mass spectrometry of enzymatically cleaved N-glycans from serum glycoproteins captured and washed on an amine-reactive slide as previously described (Figure 26).²⁷⁹ After data processing and application of data quality criteria, 55 N-glycans were detected across the samples with an average coefficient of variation (CV) of 7.9% across the technical triplicates (Table 4). The clinical descriptions of the patients included ethnicity/race, age, and BMI, and had no significant differences between the benign and cancerous samples (Table 3).

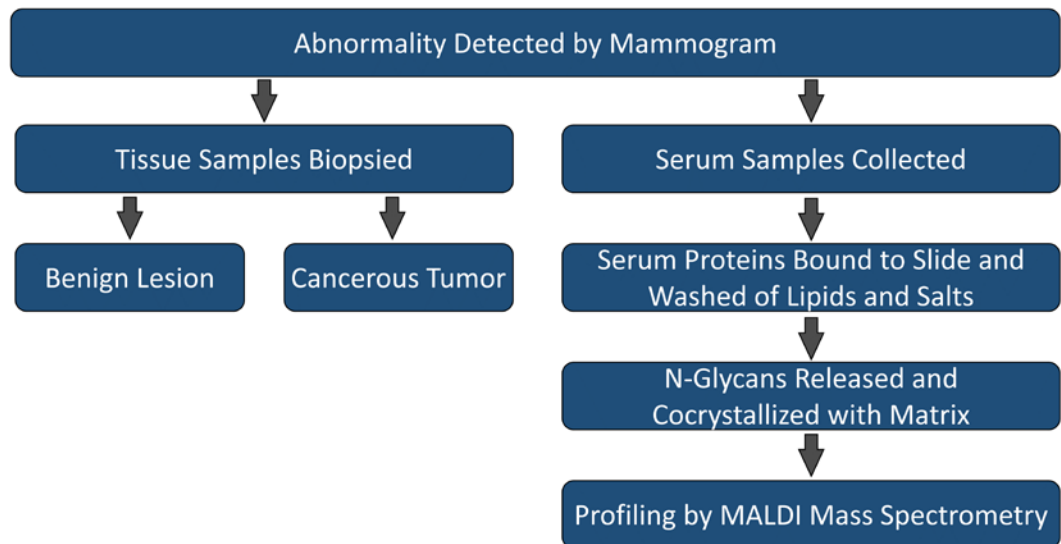


Figure 26. Sample collection, processing, and analysis overview.

Theoretical m/z	Observed m/z	Error (ppm)	Composition	Putative Structure	Oxford Nomenclature	Glycosylation Traits	Theoretical m/z	Observed m/z	Error (ppm)	Composition	Putative Structure	Oxford Nomenclature	Glycosylation Traits
1136.3964	1136.3994	2.64	Hex3HexNAc3 + Na1		A1		2157.7562	2157.7901	15.71	Hex5HexNAc5NeuAc1 + Na1		A2BG2S1	Sialylated, Bisected
1257.4226	1257.4235	0.72	Hex5HexNAc2 + Na1		M5	High Mannose	2163.7432	2163.7430	0.09	Hex4dHex1HexNAc5NeuAc1 + Na2		A2BF1G1S1	Fucosylated, Sialylated, Bisected
1282.4543	1282.4532	0.86	Hex3dHex1HexNAc3 + Na1		A1F1	Fucosylated	2174.7715	2174.7737	1.01	Hex6dHex1HexNAc5 + Na1		A3F1G3	Fucosylated, Triantennary
1298.4492	1298.4546	4.16	Hex4HexNAc3 + Na1		A1G1		2179.7382	2179.7378	0.18	Hex5HexNAc5NeuAc1 + Na2		A2BG2S1	Sialylated, Bisected
1339.4757	1339.4777	1.49	Hex3HexNAc4 + Na1		A2	Biantennary	2267.7542	2267.7533	0.38	Hex5HexNAc4NeuAc1 + Na2		A2G2S2	Sialylated, Biantennary
1419.4755	1419.4788	2.32	Hex6HexNAc2 + Na1		M6	High Mannose	2287.8192	2287.7695	21.72	Hex4dHex2HexNAc5NeuAc1 + Na1		A2BF2G1S1	Fucosylated, Sialylated, Bisected
1444.5071	1444.5094	1.59	Hex4dHex1HexNAc3 + Na1		A1G1F1	Fucosylated	2289.7361	2289.7430	3.01	Hex5HexNAc4NeuAc1 + Na3		A2G2S2	Sialylated, Biantennary
1485.5336	1485.5347	0.74	Hex3dHex1HexNAc4 + Na1		A2F1	Fucosylated, Biantennary	2303.8141	2303.8222	3.52	Hex5dHex1HexNAc5NeuAc1 + Na1		A2BF1G2S1	Fucosylated, Sialylated, Bisected
1501.5286	1501.5302	1.07	Hex4HexNAc4 + Na1		A2G1	Biantennary	2319.8090	2319.8084	0.26	Hex6HexNAc5NeuAc1 + Na1		A3G3S1	Sialylated, Triantennary
1611.5266	1611.5230	2.23	Hex4HexNAc3NeuAc1 + Na2		A1G1S1	Sialylated	2325.7961	2325.8012	2.19	Hex5dHex1HexNAc5NeuAc1 + Na2		A2BF1G2S1	Fucosylated, Sialylated, Bisected
1622.5548	1622.5512	2.22	Hex6HexNAc3 + Na1		MSA1G1	Hybrid	2341.7910	2341.7993	3.54	Hex6HexNAc5NeuAc1 + Na2		A3G3S1	Sialylated, Triantennary
1647.5865	1647.5918	3.22	Hex4dHex1HexNAc4 + Na1		A2G1F1	Fucosylated, Biantennary	2377.8509	2377.8568	2.48	Hex6dHex1HexNAc6 + Na1		A4F1G3	Fucosylated, Tetraantennary
1663.5814	1663.5847	1.98	Hex5HexNAc4 + Na1		A2G2	Biantennary	2393.8458	2393.8539	3.38	Hex7HexNAc6 + Na1		A4G4	Tetraantennary
1688.6130	1688.6168	2.25	Hex3dHex1HexNAc5 + Na1		A2BF1	Fucosylated, Bisected	2413.8121	2413.8176	2.28	Hex5dHex1HexNAc4NeuAc2 + Na2		A2F1G2S2	Fucosylated, Sialylated, Biantennary
1757.5845	1757.5832	0.74	Hex4dHex1HexNAc3NeuAc1 + Na2		A1F1G1S1	Fucosylated, Sialylated	2435.7940	2435.7983	1.77	Hex5dHex1HexNAc4NeuAc2 + Na3		A2F1G2S2	Fucosylated, Sialylated, Biantennary
1809.6393	1809.6432	2.16	Hex5dHex1HexNAc4 + Na1		A2F1G2	Fucosylated, Biantennary	2471.8540	2471.8567	1.09	Hex5dHex2HexNAc5NeuAc1 + Na2		A2BF2G2S1	Fucosylated, Sialylated, Bisected
1850.6659	1850.6736	4.16	Hex4dHex1HexNAc5 + Na1		A2BF1G1	Fucosylated, Bisected	2487.8489	2487.8517	1.13	Hex6dHex1HexNAc5NeuAc1 + Na2		A3F1G3S1	Fucosylated, Sialylated, Triantennary
1866.6608	1866.6673	3.48	Hex5HexNAc5 + Na1		A2BG2	Bisected	2522.8884	2522.9265	15.10	Hex6HexNAc6NeuAc1 + Na1		A4G3S1	Sialylated, Tetraantennary
1935.6322	1935.6335	0.67	Hex6HexNAc3NeuAc1 + Na2		MSA1G1S1	Hybrid, Sialylated	2537.8880	2537.8431	17.69	Hex5dHex2HexNAc4NeuAc2 + Na1		A2F2G2S2	Fucosylated, Sialylated, Biantennary
1938.6819	1938.6845	1.34	Hex4dHex1HexNAc4NeuAc1 + Na1		A2F1G1S1	Fucosylated, Sialylated, Biantennary	2539.9037	2539.9123	3.39	Hex7dHex1HexNAc6 + Na1		A4F1G4	Fucosylated, Tetraantennary
1954.6788	1954.6764	0.20	Hex5HexNAc4NeuAc1 + Na1		A2G2S1	Sialylated, Biantennary	2616.8915	2616.8977	2.37	Hex5dHex1HexNAc5NeuAc2 + Na2		A2BF1G2S2	Fucosylated, Sialylated, Bisected
1955.6972	1955.6833	7.11	Hex5dHex2HexNAc4 + Na1		A2F2G2	Fucosylated, Biantennary	2632.8864	2632.8964	3.80	Hex6HexNAc5NeuAc2 + Na2		A3G2S2	Sialylated, Triantennary
1960.6638	1960.6690	2.65	Hex4dHex1HexNAc4NeuAc1 + Na2		A2F1G1S1	Fucosylated, Sialylated, Biantennary	2638.8734	2638.8798	2.43	Hex5dHex1HexNAc5NeuAc2 + Na3		A2BF1G2S2	Fucosylated, Sialylated, Biantennary
1976.6588	1976.6622	1.72	Hex5HexNAc4NeuAc1 + Na2		A2G2S1	Sialylated, Biantennary	2668.9463	2668.9884	15.77	Hex6dHex1HexNAc5NeuAc1 + Na1		A4F1G3S1	Fucosylated, Sialylated, Tetraantennary
2012.7187	2012.7252	3.23	Hex5dHex1HexNAc5 + Na1		A2BF1G2	Fucosylated, Bisected	2778.9443	2778.9537	3.38	Hex6dHex1HexNAc5NeuAc2 + Na2		A3F1G3S2	Fucosylated, Sialylated, Triantennary
2028.7136	2028.7191	2.71	Hex6HexNAc5 + Na1		A3G3	Triantennary	2800.9263	2800.9384	4.32	Hex6dHex1HexNAc5NeuAc2 + Na3		A3F1G3S2	Fucosylated, Sialylated, Triantennary
2100.7347	2100.7399	2.48	Hex5dHex1HexNAc4NeuAc1 + Na1		A2F1G2S1	Fucosylated, Sialylated, Biantennary	3114.0037	3114.0199	5.20	Hex6dHex1HexNAc5NeuAc3 + Na4		A3F1G3S3	Fucosylated, Sialylated, Triantennary
2122.7167	2122.7228	2.87	Hex5dHex1HexNAc4NeuAc1 + Na2		A2F1G2S1	Fucosylated, Sialylated, Biantennary							

Table 4. N-glycans observed in serum samples of patients.

As the clinical characteristics and prognosis of DCIS and invasive breast cancer vary greatly, initial analysis was focused on detecting differences between benign, DCIS, and invasive breast cancer. No differences were detected between the N-glycan classes, with the majority of the N-glycan profile intensity being attributed to biantennary, fucosylated, and/or sialylated N-glycans (Fig. 27). As no differences in the overall N-glycan structural class profiles were detected, individual N-glycan intensities were investigated.

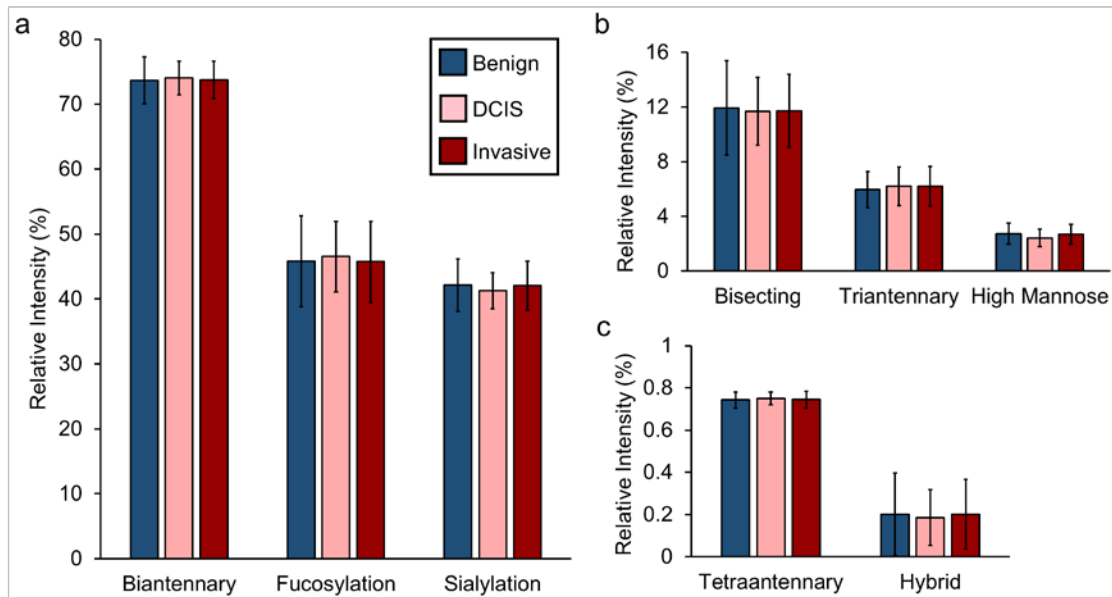


Figure 27. N-glycan class intensities of serum samples with benign lesions, DCIS, and invasive breast cancer. (A) High, (B) medium, and (C) low intensity serum N-glycan classes of benign, DCIS, and invasive samples. Data was analyzed by linear regression modeling of log-transformed relative intensities controlled for age, ethnicity/race, and BMI to identify differences across disease status. Using a 5% false-discovery rate, no significant associations were found.

There were two N-glycans (2179.738 m/z (A2BG2S1) and 2413.812 m/z (A2G2F1S2)) with significantly different intensities in the benign and DCIS serum samples (Figure 28, Table 5). The 2179.738 m/z N-glycan has a bisected and sialylated structure and had a lower intensity in the DCIS samples. The 2413.812 m/z N-glycan has a fucosylated and multiply sialylated biantennary structure, and had a higher intensity in the DCIS samples (Figure 28 a,b). Based on a logistic model using the intensities of these two N-glycans, age, ethnicity/race, and BMI, an AUROC of 0.732 (0.644 - 0.819) was found for the discrimination of benign and DCIS samples (Figure 28 c). Using the same variables, an AUROC of 0.613 (0.538 - 0.688) was found for the discrimination of benign and invasive samples (Figure 28 c). For the evaluation of the entire set based on clinical designation, there were no N-glycan classes or individual N-glycans with differences between the benign and invasive or benign and all cancerous samples.



m/z	Composition	Putative Structure	Oxford Nomenclature	Intensity in In Situ	Fold Change
2179.7382	Hex5HexNAc5NeuAc1 + Na2		A2BG2S1	Lower	0.826
2413.8121	Hex5dHex1HexNAc4NeuAc2 + Na2		A2G2F1S2	Higher	1.201

Table 5. Serum N-glycans with significant differences between benign and DCIS patients.

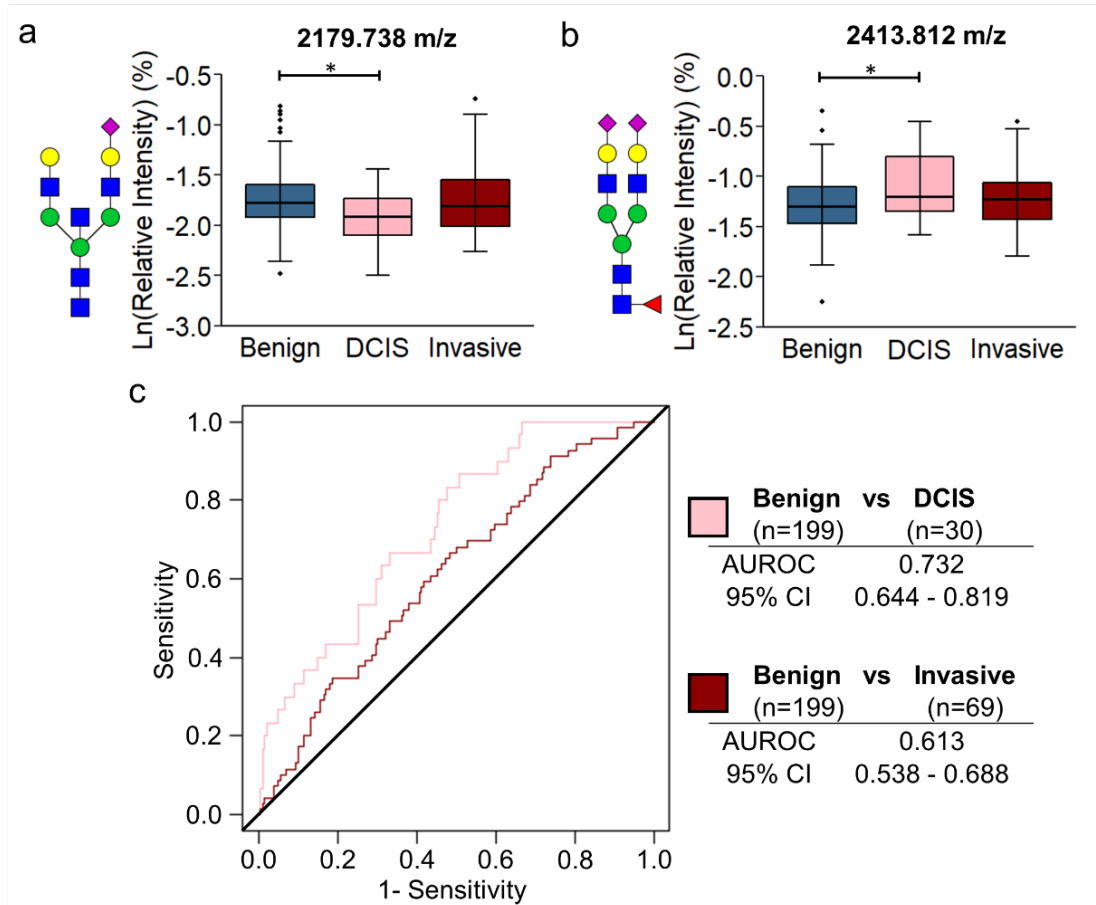


Figure 28. Serum N-glycan differences between benign, DCIS, and invasive samples. Log-transformed relative intensity values of the (A) 2179.738 m/z and (B) 2413.812 m/z N-glycans had significant differences in benign and DCIS samples using a 5% false discovery rate. (C) Using the log-transformed relative intensities of these N-glycans, age, BMI, and ethnicity/race, benign and DCIS samples could be distinguished with an AUROC of 0.732, and benign and invasive samples could be distinguished with an AUROC of 0.613.

Additional analysis identified significant relationships between N-glycan intensities and patient age or BMI. There were 24 N-glycans associated with age, as well as the triantennary, tetraantennary, and fucosylation N-glycan classes (Table 6). Additionally, there were 12 N-glycans associated with BMI, as well as the biantennary, triantennary, and tetraantennary N-glycan classes (Table 7). While some of these associations were present in both benign and cancerous samples (Figure 29 a,c), others were only present in benign samples or displayed differences between the samples at certain ranges (Figure 29 b,d). These findings prompted an in-depth analysis of the samples stratified by age and BMI.

m/z	Composition	Putative Structure	Oxford Nomenclature	Intensity Change as Age Increases	m/z	Composition	Putative Structure	Oxford Nomenclature	Intensity Change as Age Increases
1136.3964	Hex3HexNAc3 + Na1		A1	Increases	2122.7167	Hex5dHex1HexNAc4NeuAc1 + Na2		A2F1G2S1	Decreases
1339.4757	Hex3HexNAc4 + Na1		A2	Increases	2163.7432	Hex4dHex1HexNAc5NeuAc1 + Na2		A2BF1G1S1	Increases
1444.5071	Hex4dHex1HexNAc3 + Na1		A1G1F1	Decreases	2267.7542	Hex5HexNAc4NeuAc2 + Na2		A2G2S2	Increases
1485.5336	Hex3dHex1HexNAc4 + Na1		A2F1	Increases	2319.8090	Hex6HexNAc5NeuAc1 + Na1		A3G3S1	Increases
1501.5286	Hex4HexNAc4 + Na1		A2G1	Increases	2341.7910	Hex6HexNAc5NeuAc1 + Na2		A3G3S1	Increases
1647.5865	Hex4dHex1HexNAc4 + Na1		A2G1F1	Decreases	2377.8509	Hex6dHex1HexNAc6 + Na1		A4F1G3	Increases
1688.6130	Hex3dHex1HexNAc5 + Na1		A2BF1	Increases	2471.8540	Hex5dHex2HexNAc5NeuAc1 + Na2		A2BF2G2S1	Decreases
1757.5845	Hex4dHex1HexNAc3NeuAc1 + Na2		A1F1G1S1	Decreases	2522.8884	Hex6HexNAc8NeuAc1 + Na1		A4G3S1	Increases
1809.6393	Hex5dHex1HexNAc4 + Na1		A2F1G2	Decreases	2539.9037	Hex7dHex1HexNAc6 + Na1		A4F1G4	Increases
1850.6659	Hex4dHex1HexNAc5 + Na1		A2BF1G1	Increases	2632.8864	Hex6HexNAc5NeuAc2 + Na2		A3G2S2	Increases
1866.6608	Hex5HexNAc5 + Na1		A2BG2	Increases	Fucosylation	-	-	-	Decreases
1955.6972	Hex5dHex2HexNAc4 + Na1		A2F2G2	Increases	Triantennary	-	-	-	Increases
2012.7187	Hex5dHex1HexNAc5 + Na1		A2BF1G2	Decreases	Tetraantennary	-	-	-	Increases
2100.7347	Hex5dHex1HexNAc4NeuAc1 + Na1		A2F1G2S1	Decreases					

Table 6. Serum N-glycans with significant associations with age.

m/z	Composition	Putative Structure	Oxford Nomenclature	Intensity Change as BMI Increases	m/z	Composition	Putative Structure	Oxford Nomenclature	Intensity Change as BMI Increases
1501.5286	Hex4HexNAc4 + Na1		A2G1	Increases	2668.9463	Hex6dHex1HexNAc5NeuAc1 + Na2		A3F1G3S1	Increases
1809.6393	Hex5dHex1HexNAc4 + Na1		A2F1G2	Decreases	2778.9443	Hex6dHex1HexNAc6NeuAc1 + Na1		A4F1G3S1	Increases
2163.7432	Hex4dHex1HexNAc5NeuAc1 + Na2		A2BF1G1S1	Increases	2800.9263	Hex6dHex1HexNAc5NeuAc2 + Na2		A3F1G3S2	Increases
2303.8141	Hex4dHex1HexNAc5NeuAc1 + Na2		A2BF1G1S1	Increases	3114.0037	Hex6dHex1HexNAc5NeuAc2 + Na3		A3F1G3S2	Increases
2413.8121	Hex7HexNAc6 + Na1		A4G4	Increases	Biantennary	Hex6dHex1HexNAc5NeuAc3 + Na4		A3F1G3S3	Decreases
2435.7940	Hex5dHex1HexNAc4NeuAc2 + Na2		A2F1G2S2	Increases	Triantennary	-	-	-	Increases
2487.8489	Hex5dHex1HexNAc4NeuAc2 + Na3		A2F1G2S2	Increases	Tetraantennary	-	-	-	Increases

Table 7. Serum N-glycans with significant associations with BMI.

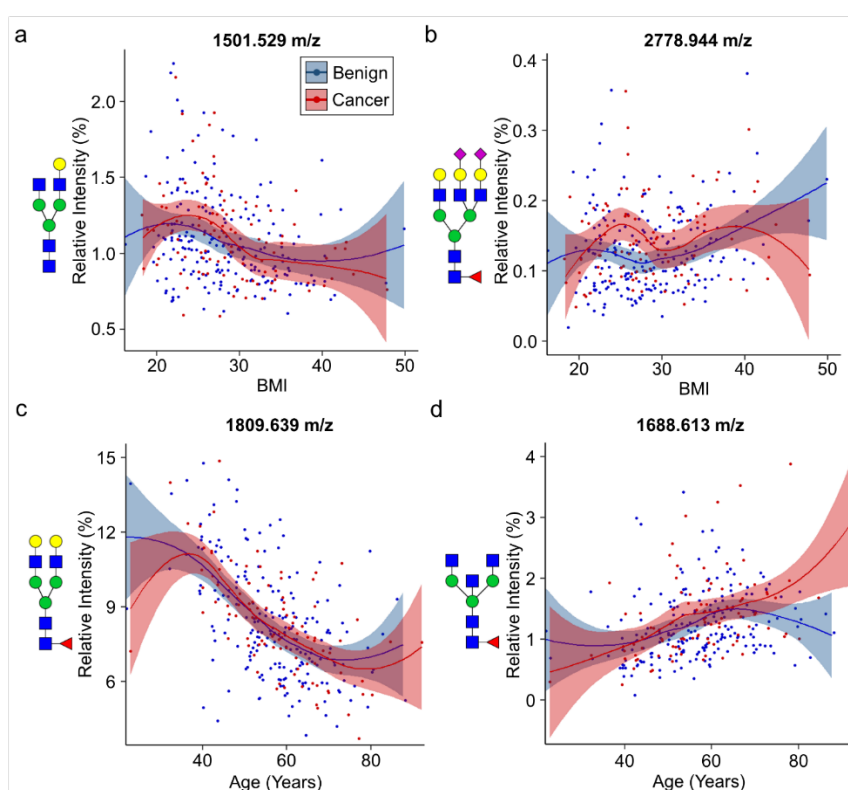


Figure 29. Correlation of serum N-glycan intensities with the patient age and BMI. The patient BMI and the intensities of (A) 1501.529 m/z and (B) 2778.944 m/z N-glycans. The patient age and the intensities of both (C) 1809.639 m/z and (D) 1688.613 m/z N-glycans were significantly associated.

The samples were separated by age (less than 40, 40 - 49, 50 - 74, and older than 74) and/or by BMI (less than 18.5, 18.5 - 24.9, 25 - 30, and more than 30). There were no significant findings between the benign and cancerous samples in any individual or combination of these groups, except for the patients that were 50 - 74 years old and within 18.5 - 24.9 BMI. These samples had significant differences in intensity between benign (n = 45) and cancerous samples (n = 11) for two N-glycans with, 1850.666 m/z (A2BG1F1) and 2163.743 m/z (A2BG1F1S1), and the bisecting N-glycan class (Figure 30 a-c, Table 8). Based on a logistic model using the intensities of 1850.666 m/z and 2163.743 m/z, age, and BMI an AUROC of 0.899 (0.801 - 0.997) found for the discrimination of benign and cancerous samples with an 82% sensitivity and 84% specificity (Figure 30 d).



m/z	Composition	Putative Structure	Oxford Nomenclature	Intensity in Cancer	Fold Change
1850.6659	Hex4dHex1HexNAc5 + Na1		A2BG1F1	Lower	0.726
2163.7432	Hex4dHex1HexNAc5NeuAc1 + Na2		A2BG1F1S1	Lower	0.733
Bisecting	-	-	-	Lower	0.772

Table 8. Serum N-glycans with significant differences between benign and cancer patients of age 50 - 74 and BMI 18.5 - 24.9.

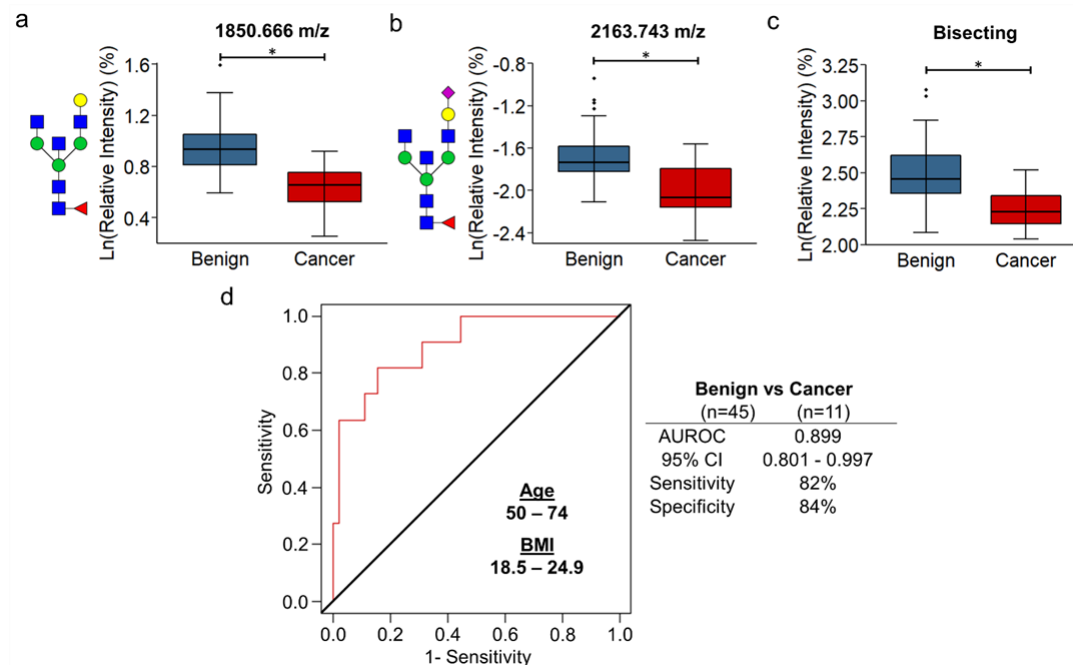


Figure 30. Serum N-glycan differences between benign and cancer samples from patients aged 50 – 74 and with a BMI of 18.5 – 24.9. Log-transformed relative intensity values of the (A) 1850.666 m/z and (B) 2163.743 m/z N-glycans and (C) bisecting N-glycan class had significant differences in benign and cancer samples using a 5% false discovery rate. (D) Using the log-transformed relative intensities of the two individual N-glycans, age, and BMI, benign and cancer samples were classified using the AUROC of the logistic regression model.

For the IgG analysis of the samples, 34 N-glycans were reliably and consistently detected across all the samples (Table 9). Across all samples, the overall IgG N-glycan profile was mostly composed of biantennary and fucosylated N-glycans (Figure 31). High mannose, hybrid, and triantennary N-glycans comprised less than 1% of the IgG N-glycan profile across the benign, DCIS, and invasive samples. There were no significant differences detected in the individual N-glycans or N-glycan classes between the benign samples and the DCIS or invasive samples, individually or combined. The samples were stratified into the same groups by BMI and age as described for the serum sample analysis, and there were no statistically significant differences in these groups either.

Theoretical m/z	Observed m/z	Error (ppm)	Composition	Putative Structure	Oxford Nomenclature	Glycosylation Traits	Theoretical m/z	Observed m/z	Error (ppm)	Composition	Putative Structure	Oxford Nomenclature	Glycosylation Traits
1136.3964	1136.3989	2.20	Hex3HexNAc3 + Na1		A1		1809.6393	1809.6387	0.33	Hex5dHex1HexNAc4 + Na1		A2F1G2	Fucosylated, Biantennary
1257.4226	1257.4287	4.85	Hex5HexNAc2 + Na1		M5	High Mannose	1850.6659	1850.6689	1.62	Hex4dHex1HexNAc5 + Na1		A2BF1G1	Fucosylated, Bisected
1282.4543	1282.4584	3.20	Hex3dHex1HexNAc3 + Na1		A1F1	Fucosylated	1866.6608	1866.6625	0.91	Hex5HexNAc5 + Na1		A2G2	Bisected
1298.4492	1298.4533	3.16	Hex4HexNAc3 + Na1		A1G1		1954.6768	1954.6807	2.00	Hex5HexNAc4NeuAc1 + Na1		A2G2S1	Sialylated, Biantennary
1339.4757	1339.4762	0.37	Hex3HexNAc4 + Na1		A2	Biantennary	1960.6638	1960.6712	3.77	Hex4dHex1HexNAc4NeuAc1 + Na2		A2F1G1S1	Fucosylated, Sialylated, Biantennary
1419.4755	1419.4768	0.92	Hex6HexNAc2 + Na1		M6	High Mannose	1976.6588	1976.6608	1.01	Hex5HexNAc4NeuAc1 + Na2		A2G2S1	Sialylated, Biantennary
1444.5071	1444.5073	0.14	Hex4dHex1HexNAc3 + Na1		A1G1F1	Fucosylated	2012.7187	2012.7194	0.35	Hex5dHex1HexNAc5 + Na1		A2BF1G2	Fucosylated, Bisected
1460.5020	1460.5064	3.01	Hex5HexNAc3 + Na1		M4A1G1	Hybrid	2100.7347	2100.7453	5.05	Hex5dHex1HexNAc4NeuAc1 + Na1		A2F1G2S1	Fucosylated, Sialylated, Biantennary
1485.5336	1485.5397	4.11	Hex3dHex1HexNAc4 + Na1		A2F1	Fucosylated, Biantennary	2122.7167	2122.7268	4.76	Hex5dHex1HexNAc4NeuAc1 + Na2		A2F1G2S1	Fucosylated, Sialylated, Biantennary
1501.5286	1501.5278	0.53	Hex4HexNAc4 + Na1		A2G1	Biantennary	2157.7562	2157.7725	7.55	Hex5HexNAc5NeuAc1 + Na1		A2B2S1	Sialylated, Bisected
1542.5551	1542.5594	2.79	Hex3HexNAc5 + Na1		A2B	Bisect	2163.7432	2163.7686	11.74	Hex4dHex1HexNAc5NeuAc1 + Na2		A2BF1G1S1	Fucosylated, Sialylated, Bisected
1581.5283	1581.5302	1.20	Hex7HexNAc2 + Na1		M7	High Mannose	2174.7715	2174.7895	8.28	Hex6dHex1HexNAc5 + Na1		A3F1G3	Fucosylated, Triantennary
1622.5548	1622.5561	0.80	Hex6HexNAc3 + Na1		M5A1G1	Hybrid	2179.7382	2179.7612	10.55	Hex5HexNAc5NeuAc1 + Na2		A2B2S1	Sialylated, Bisected
1647.5865	1647.5885	1.21	Hex4dHex1HexNAc4 + Na1		A2G1F1	Fucosylated, Biantennary	2303.8141	2303.8216	3.26	Hex5dHex1HexNAc5NeuAc1 + Na1		A2BF1G2S1	Fucosylated, Sialylated, Bisected
1663.5814	1663.5812	0.12	Hex5HexNAc4 + Na1		A2G2	Biantennary	2319.8090	2319.8177	3.75	Hex6HexNAc5NeuAc1 + Na1		A3G3S1	Sialylated, Triantennary
1688.6130	1688.6131	0.06	Hex3dHex1HexNAc5 + Na1		A2BF1	Fucosylated, Bisected	2325.7961	2325.8098	5.89	Hex5dHex1HexNAc5NeuAc1 + Na2		A2BF1G2S1	Fucosylated, Sialylated, Bisected
1704.6080	1704.6120	2.35	Hex4HexNAc5 + Na1		A2BG1	Bisect	2341.7910	2341.8061	6.45	Hex6HexNAc5NeuAc1 + Na2		A3G3S1	Sialylated, Triantennary

Table 9. N-glycans observed in IgG from serum samples of patients.

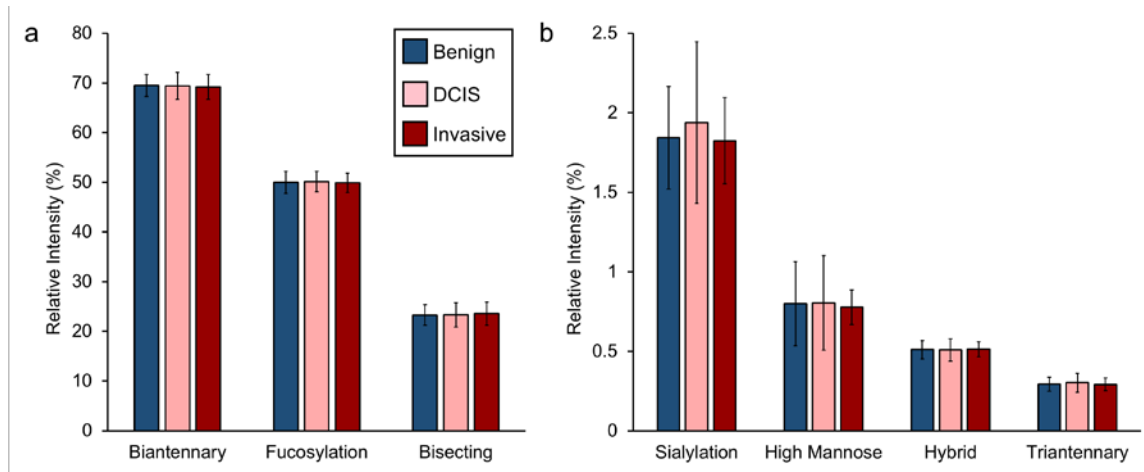


Figure 31. N-glycan class intensities of IgG from serum samples with benign lesions, DCIS, and invasive breast cancer. (A) High and (B) low intensity serum N-glycan classes of benign, DCIS, and invasive samples. Data was analyzed by linear regression modeling of log-transformed relative intensities controlled for age, ethnicity/race, and BMI to identify differences across disease status. Using a 5% false-discovery rate, no significant associations were found.

5.5 Discussion

In this study, the serum and IgG N-glycan profiles of patients with an abnormality detected by a mammogram were compared to evaluate their ability to distinguish benign lesions from breast cancer. The samples were processed using recently developed high-throughput MALDI imaging mass spectrometry workflows.^{20,279} When analyzing the serum N-glycan profiles of all the samples, two N-glycans had significant differences in the benign and DCIS samples. Incorporating the intensities of these N-glycans with age, ethnicity/race, and BMI had moderate diagnostic potential for DCIS samples but not invasive samples.

There were many N-glycan intensities that had altered associations with age or BMI, depending on the diagnosis of the patient. Stratifying the samples by age and BMI identified two N-glycans and two N-glycan classes with significant intensity differences in the benign and cancerous samples of patients aged 50 - 74 with a BMI of 18.5 - 24.9, and when combined with age and BMI had high diagnostic ability for this patient subset.

Inflammatory stimuli due to the presence of a disease can cause an alteration in the glycosylation of plasma cells and hepatocytes, which produce the bulk of serum proteins.^{304,305} Altered serum N-glycan profiles have been associated with a range of diseases and conditions, including ulcerative colitis, diabetes, Crohn's disease, gastric cancer, ovarian cancer, Alzheimer's disease, hepatocellular carcinoma, and multiple myeloma.^{96,101,104,306-310} Serum analysis of breast cancer patients have identified N-glycan changes associated with circulating tumor cell counts, breast cancer subtypes, prognosis, breast cancer progression, and lymph node metastasis.^{206,207,210,217,219,229} While several studies have investigated breast cancer detection by serum N-glycan analysis, they have either had healthy women donors as controls, used methods unable to identify individual N-glycans, or only analyzed certain types of N-glycans.^{102,206,216,218,220} To properly address the issue of false-positives in mammograms, the samples from patients with breast cancer need to be compared to samples from patients with a lesion detected by a mammogram that requires a biopsy, because healthy women with no abnormalities detected by mammogram do not suffer from the issues of false-positives. This cohort can appropriately investigate the potential of

serum N-glycan analysis for breast cancer detection because of the clinical relevance of the benign samples, the controlled surgical collection site, and the availability and incorporation of age, ethnicity/race, and BMI into the analysis.

Age and BMI have strong associations with breast cancer risk, as well as specific alterations in the serum N-glycan profile.^{81–83,155,311–314} In this study, there were 24 N-glycans with significant associations with age, and 12 N-glycans with significant associations with BMI (Table 6 and Table 7). Age and BMI were controlled for in the initial analysis to account for these associations and isolate N-glycan intensity differences that were due to the presence of breast cancer. Further investigation found that only samples within certain ranges of age and BMI had disease-associated serum N-glycan changes (Figure 29 b,d). To identify any disease-associated serum N-glycan intensities changes only present in a subset of the cohort, the samples were stratified according to age and BMI. To remain clinically relevant, the samples were grouped based on the USPSTF mammogram recommendation for age and the BMI classifications for underweight, healthy weight, overweight, and obese.²⁹⁸

The subset of samples from patients with a BMI of 18.5 - 24.9 and aged 50 - 74 could be readily distinguished as having a benign lesion or breast cancer using the intensities of two N-glycans, age, and BMI (Figure 30 d). While this is a subset of the population, it makes up a significant portion of the U.S. women that will have mammogram screening for breast cancer. For women with an average-risk of breast cancer, the USPSTF recommends only women aged 50 - 74 to have biennial mammography screening and make up the bulk of women

that will have a mammogram.²⁹⁸ Results from the 2017 - 2018 National Health and Nutrition Examination Survey (NHANES) indicated that approximately one third of U.S. adult women are within the 18.5 - 24.9 BMI range.³¹⁵ The healthy weight women may have had the best N-glycan biomarker performance due to the lack of systemic inflammation that is present in overweight and obese women.³¹⁶ The plasma cells and hepatocytes, from which the majority of serum glycoproteins originate, undergo significant cellular changes in presence of inflammatory stimuli. The serum N-glycan changes found in cancer patients may be masked, altered, or negated by the N-glycan changes due to the obesity-induced inflammatory stimuli.

There were several limitations to this study. Breast cancer is an incredibly heterogeneous disease, that includes multiple molecular and genomic subtypes with distinct clinical behavior and biological characteristics: luminal-A, luminal-B, HER2-positive, and basal-like. The risk factors for breast cancer, including age, body mass index (BMI), onset of menopause, parity, family-history, and personal history of benign breast disease differ in effect for each molecular subtype.¹⁶⁴ As subtype information was not available for this cohort, there may be N-glycan alterations only in certain subtypes that are not being detected. Additionally, there were fewer samples from Asian or African American patients. This makes it more difficult to account for variances due to ethnicity/race, especially after stratifying the cohort into smaller subsets.

The inability to identify any IgG N-glycan profile alterations across the sample groups, stratified or not, was unexpected for several reasons. Many of

the distinctive N-glycans from the serum N-glycan analysis make up a significant portion of the typical IgG N-glycan profile, and IgG contributes a large portion of serum's protein abundance.³¹⁷ One possible explanation for the lack of altered IgG N-glycans or classes could be that, as described above, the inflammatory stimuli from other clinical factors were masking the breast cancer-associated N-glycan alterations, and/or the methodology was not robust or sensitive enough to detect the alterations in the smaller sample subsets. The workflow used in this study did not account for the contribution of the antibody's N-glycans to the samples. While hypothetically this should be equal across all samples, the small volumes used for the antibody spotting are more susceptible to differences in total antibody amount spotted. Future studies utilizing this IgG N-glycan analysis workflow should test if deglycosylating the antibody spots prior to adding the samples affects target protein capture efficiency and specificity.

Additional studies identifying IgG or other protein(s) that could have distinctive N-glycans could increase the performance of the N-glycans as breast cancer detection biomarkers. Recent glycoproteomic approaches have made high-throughput analysis of complex biological fluids quicker and more sensitive.³¹⁸ Glycopeptide analysis of this cohort could identify the proteins that contain the cancer-associated N-glycans identified in this study. Isolating low-abundant proteins shed from breast cancer cells or released from dying cancer cells could identify N-glycan changes with a higher specificity across all age and BMI ranges. To further investigate the fucosylated bisecting N-glycans that had a significant association with breast cancer, the location of the fucose could be

determined by applying Endoglycosidase F3 to the samples instead of PNGase F to detect only core fucosylated N-glycans.²⁷⁹

Chapter 6: Conclusions, Limitations, and Future Studies

6.1 Overall Findings

As the understanding of the role of N-glycosylation in disease and immune recognition develops, there becomes a greater need for high-throughput assays capable of analyzing cohorts with hundreds or thousands of samples. The work presented in this dissertation addresses this need by describing the development and validation of a novel method for N-glycan profiling of serum and plasma (Chapter 3), as well as urine and prostatic fluids (Chapter 4). The utility of this biofluid N-glycan profiling was demonstrated by the identification of serum N-glycan differences between women with benign masses or breast cancer that could aid breast cancer screening (Chapter 5). The conclusions, limitations, and future research opportunities of each of these works are discussed in greater detail below.

6.2 Serum and Plasma N-glycan Profiling by MALDI-IMS

6.2.1 Conclusions

The development of this novel method was based on traditional histology and mass spectrometry imaging workflows. Current high-throughput N-glycan profiling methods overcome the noise and ionization suppression caused by the complex and abundant array of biomolecules present in serum and plasma using lengthy derivatizations and clean up steps using chromatography. This method was able to circumvent the need for these steps by using a slide-based assay. The key step in the workflow is the binding of the serum glycoproteins to an amine-reactive slide with a covalent bond. This allowed for the samples to be

quickly washed to remove the lipids and salts that would reduce or completely diminish the ionization of the enzymatically released N-glycans.

Being a slide-based assay, this method utilizes the well-established and optimized protocols for tissue N-glycan profiling to achieve a high sensitivity and data acquisition speed. Using an enzyme to release the N-glycans from the serum or plasma glycoproteins provides flexibility to the method. PNGase F can be used to analyze all N-glycans, while using Endo F3 will target only the core fucosylated N-glycans. Endoglycosidases can be paired with exoglycosidases to target specific N-glycan structures. Doing so with sialidase and PNGase F increased the sensitivity and number of low abundant polylectosamine structures detected. The simple processing, adaptability, and high-throughput of this method will facilitate large-scale biomarker detection studies using patient serum or plasma samples.

6.2.2 Limitations and Future Studies

The method described in this chapter achieves short processing times by avoiding derivatizing typically needed for extracting or increasing the ionization efficiency of the released N-glycans, but one of the other benefits of derivatization is that a linkage-specific addition of chemical groups can be incorporated. Due to the difference in mass of the chemical groups, this allows for identification of N-glycan structural isomers by calculating the mass shift. The non-MS based high-throughput serum N-glycan profiling methods can separate isomers using either HILIC columns for UHPLC or the different migration speeds

through the polymer-filled capillary under an electric field for xCE. With the output of this method being the m/z and no derivatization steps included, this method cannot identify N-glycan isomers. Future research will focus on overcoming this by performing an amidation-oxidation derivatization, as has been described and optimized for tissue MSI N-glycan analysis, after the samples have been covalently bound to the slide.³¹⁹ The inclusion of this processing step may require optimization of reaction components, concentrations, and lengths. An alternative approach for gaining structural information would be to sequentially digest the N-glycans with varying glycosidases as has been done for UHPLC analysis. The benefits of structural isomer identification would come at the cost of increased processing times. These factors would need to be weighed when designing the experimental approach.

In its current format, the diluting, spotting, and washing steps of this method require hands-on time that can be laborious. While this has been partially mitigated with the use of an electronic pipette with a multi-dispense feature for sample spotting and a multichannel pipette for the wash steps, these steps could be automated with a liquid handling robot. This would not only decrease the hands-on time, but it would also increase the precision of the workflow and make the processing even more simple than it already is. All of these factors are important for developing a workflow that could be implemented in a clinical setting as a diagnostic assay.

While this method has currently been used for N-glycan analysis, other biomolecules could be targeted. This can be done by switching the enzyme that

is applied to the samples to either chondroitinase, collagenase, or trypsin for glycosaminoglycan and proteomic analysis. Effective manners of applying these enzymes and analyzing their products have already been established for MALDI analysis. These enzymes could be incorporated into a multiplexed workflow for the analysis of biofluid samples in a manner similar to what has been established for tissue analysis by MALDI mass spectrometry.³²⁰ Once the sample is bound to the slide, it can be analyzed with one enzyme, the matrix can be washed away, and then another enzyme could be used.

6.3 Urine and Prostatic Fluid N-glycan Profiling by MALDI-IMS

6.3.1 Conclusions

In chapter 2, an N-glycan profiling workflow applicable to urine and prostatic fluids was presented. This workflow was adapted from the serum and plasma N-glycan profiling workflow discussed in chapter 1 and only required an additional filtering step using a molecular weight filter and a benchtop centrifuge. Like the serum and plasma method, this method can also be used with other enzymes besides PNGase F, like endoglycosidase F3 specific to core fucoses,²⁴⁹ or other glycosidases specific to other N-glycan structural classes. The method was developed using pooled urine samples from either healthy male or healthy female individuals. The male and female urine N-glycan profiles looked very similar but contained a larger number of sulfated N-glycans than typically seen in serum or plasma. Correspondingly, prostatic fluids collected directly had a much

lower intensity of sulfated N-glycans when compared to prostatic fluids voided in urine.

The investigation of prostatic fluids is a promising approach for biomarker discovery studies of prostatic diseases, because of the presence of shed cells and secreted proteins either directly from the diseased tissue or from surrounding tissue. Interestingly, MSI analysis of N-glycans in the luminal compartments of prostate tissue glands found a significant overlap in N-glycans detected in prostatic fluids, confirming the validity of this approach. The sensitivity of this method allowed for the detection of N-glycans that have large differences in intensities. This is a promising finding as the disease-associated N-glycan alterations may cause minor changes in the overall N-glycan profile.

6.3.2 Limitations and Future Studies

Just as the serum and plasma method was adapted, nearly each step of this workflow can also be modified for targeted analysis of specific N-glycans, proteins, or biofluids. The use of the concentration and solvent/buffer exchange step in the workflow can be adapted to other biofluids with protein concentrations similar to urine, i.e., saliva, cerebrospinal fluid, bronchial lavage. Additionally, instead of a molecular weight filter, an affinity column could be used to extract abundant proteins for analysis of low-abundant proteins. This could specifically be helpful for removing albumin as it is one of the most abundant proteins in urine and is not N-glycosylated. Alternatively, uromodulin, the most abundant

protein in urine, could be extracted either for direct analysis of its N-glycan profile or to allow for higher sensitivity of other lower concentration urine glycoproteins.

The filtration step can be modified to change the protein concentrations of the samples prior to spotting. In this way, the sensitivity of the method to low-abundant N-glycans in the starting material can be increased by increasing the amount of starting material. One of the limitations of this study is that there can be wide variation in the protein concentration of urine samples. It is very difficult to obtain absolute protein concentration measurements and adding this process to the workflow would significantly add to the labor and processing time. To avoid this, the intensities of the N-glycans are converted into relative intensities to account for differences in protein concentrations. Future development of this method could incorporate an internal standard in the form of a protein whose N-glycans can be distinguished by a mass shift introduced through a methylamidation derivatization.

6.4 Classifying Suspicious Mammogram Findings by Serum and IgG N-glycan Profiling

6.4.1 Conclusions

In this chapter, a MALDI imaging mass spectrometry approach was utilized to detect serum and IgG N-glycans and N-glycan classes that could distinguish patients with an abnormal mammogram result as having a benign lesion or breast cancer. For a significant portion of the women screened by mammogram, fucosylated bisecting N-glycans could readily distinguish patient

groups. A surprising result was that no IgG N-glycans or N-glycan classes could be used to distinguish patients that had benign lesions, DCIS, or invasive breast cancer. Even when stratifying the population based on age and BMI, there were no significant differences.

While many studies have identified N-glycan differences in healthy women compared to women with breast cancer, this study is one of the first to address the issue of breast cancer screening false-positives. The samples were collected from women at a surgical clinic after a suspicious mammogram finding making this cohort ideal for investigating this unmet clinical need. Using samples collected at a single site reduced the impact of confounding factors such as inconsistent collection or storage methods. Incorporating clinical factors that have well-established associations with breast cancer with certain N-glycans led to a more robust classification of samples. This highlights the importance of collecting and utilizing biological characteristics when investigating serum N-glycan profiles.

6.4.2 Limitations and Future Studies

One of the factors that was incorporated into the initial analysis of all the samples was ethnicity/race. There were greater than 5 samples for each ethnicity/race (Caucasian, Hispanic, African American, and Asian), but when the samples were stratified by BMI and age, there were some groupings that had no samples from either African American or Asian women. As this was the case, ethnicity/race was not included in the logistic regression model used to classify samples as being from a woman with a benign lesion or breast cancer. This does

limit the conclusions that can be made from this data. Compared to other ethnicity and racial groups, African-American women have a higher mortality rate, are diagnosed at a younger age, and are more likely to have unfavorable tumor characteristics, such as having triple-negative breast cancer.³²¹ Future studies should look to increase sample size and distributions of different ethnicity/races to not only broaden conclusions, but increase the likelihood of identifying N-glycan alterations that may be ethnicity/race-specific.

While the ideal biomarker can be universally applied, the strongest use of N-glycans as classifiers of disease status was in a subset of women aged 50-74 that were within the healthy weight range. This may be due to breast cancer-associated alterations in the N-glycan profile being masked by the significant effects of age and BMI-induced inflammation on N-glycosylation patterns of highly abundant serum proteins, such as IgG. The lack of significant alterations in the IgG N-glycan profile, indicates that the differentially expressed N-glycans in the serum are most likely carried by proteins originating from the liver. Future research should focus on using antibody-panel based N-glycan profiling of these proteins.

6.5 Final Thoughts

Overall, this dissertation has described the development of a novel biofluid N-glycan profiling technique and its application for the identification of N-glycan alterations that can be used to distinguish women with benign lesions from women with breast cancer. The workflow described in chapter 3 has the potential to become a platform for N-glycan analysis of various biofluids. This was validated in chapter 4 when the workflow was adapted for urine and prostatic fluids. The utility and clinical relevance of this approach was demonstrated by its application to breast cancer biomarker detection in chapter 5. The general workflow described in this dissertation constitutes a novel MALDI analysis platform for N-glycan profiling of multiple biofluids that has the potential for multiplexed analysis of other biomolecules. This work represents another advancement in MALDI mass spectrometry, which itself is quickly becoming a technology that can be used as a comprehensive strategy for clinical research by analysis of tissues, cells, targeted proteins, and biofluids.

REFERENCES

1. Niessen, W. M. A. & Falck, D. Introduction to Mass Spectrometry, a Tutorial. in *Analyzing Biomolecular Interactions by Mass Spectrometry* (eds. Kool, J. & Niessen, W. M. A.) 1–54 (John Wiley & Sons, Ltd, 2013). doi:10.1002/9783527673391.CH1
2. Lewis, J. K., Wei, J. & Siuzdak, G. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry in Peptide and Protein Analysis. in *Encyclopedia of Analytical Chemistry* 5880–5894 (2006). doi:10.1002/9780470027318.a1621
3. Dreisewerd, K. Recent methodological advances in MALDI mass spectrometry. *Anal. Bioanal. Chem.* **406**, 2261–2278 (2014).
4. Karas, M., Glückmann, M. & Schäfer, J. Ionization in matrix-assisted laser desorption/ionization: Singly charged molecular ions are the lucky survivors. *J. Mass Spectrom.* **35**, 1–12 (2000).
5. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science (80-.)*. **312**, 212–217 (2006).
6. Calvano, C. D., Monopoli, A., Cataldi, T. R. I. & Palmisano, F. MALDI matrices for low molecular weight compounds: an endless story? *Anal. Bioanal. Chem.* **410**, 4015–4038 (2018).
7. Wäldchen, F., Mohr, F., Wagner, A. H. & Heiles, S. Multifunctional Reactive MALDI Matrix Enabling High-Lateral Resolution Dual Polarity MS Imaging and Lipid C=C Position-Resolved MS² Imaging. *Anal. Chem.* **92**, 14130–14138 (2020).
8. Towers, M. W. & Cramer, R. Ionic liquids and other liquid matrices for sensitive MALDI MS analysis. *Adv. MALDI Laser-Induced Soft Ioniz. Mass Spectrom.* 51–64 (2015). doi:10.1007/978-3-319-04819-2_3
9. Abdelhamid, H. N. Ionic Liquids Matrices for Laser Assisted Desorption/Ionization Mass Spectrometry. *Mass Spectrom. Purif. Tech.* **01**, 109 (2015).
10. Kostrzewa, M. Application of the MALDI Biotyper to clinical microbiology: progress and potential. *Expert Rev. Proteomics* **15**, 193–202 (2018).
11. Mellmann, A. *et al.* High interlaboratory reproducibility of matrix-assisted laser desorption ionization-time of flight mass spectrometry-based species identification of nonfermenting bacteria. *J. Clin. Microbiol.* **47**, 3732–3734 (2009).
12. Drake, R. R., Boggs, S. R. & Drake, S. K. Pathogen identification using mass spectrometry in the clinical microbiology laboratory. *J. Mass Spectrom.* **46**, 1223–1232 (2011).
13. Schwamborn, K. & Caprioli, R. M. Molecular imaging by mass spectrometry-looking beyond classical histology. *Nature Reviews Cancer* **10**, 639–646 (2010).
14. Angel, P. M. & Caprioli, R. M. Matrix-assisted laser desorption ionization imaging mass spectrometry: In situ molecular mapping. *Biochemistry* **52**, 3818–3828 (2013).

15. Porta Siegel, T. *et al.* Mass Spectrometry Imaging and Integration with Other Imaging Modalities for Greater Molecular Understanding of Biological Tissues. *Molecular Imaging and Biology* **20**, 888–901 (2018).
16. Deutskens, F., Yang, J. & Caprioli, R. M. High spatial resolution imaging mass spectrometry and classical histology on a single tissue section. *J. Mass Spectrom.* **46**, 568–571 (2011).
17. Drake, R. R. *et al.* Defining the human kidney N-glycome in normal and cancer tissues using MALDI imaging mass spectrometry. *J. Mass Spectrom.* **55**, e4490 (2020).
18. McDowell, C. T., Lu, X., Mehta, A. S., Angel, P. M. & Drake, R. R. Applications and continued evolution of glycan imaging mass spectrometry. *Mass Spectrom. Rev.* 1–32 (2021).
19. Angel, P. M. *et al.* A Rapid Array-Based Approach to N-Glycan Profiling of Cultured Cells. *J. Proteome Res.* **18**, 3630–3639 (2019).
20. Black, A. P., Angel, P. M., Drake, R. R. & Mehta, A. S. Antibody Panel Based N-Glycan Imaging for N-Glycoprotein Biomarker Discovery. *Curr. Protoc. Protein Sci.* **98**, (2019).
21. Bruker Daltonics Inc. *SolariX User Manual.* (2017).
22. Silveira, J. A., Michelmann, K., Ridgeway, M. E. & Park, M. A. Fundamentals of Trapped Ion Mobility Spectrometry Part II: Fluid Dynamics. *J. Am. Soc. Mass Spectrom.* **27**, 585–595 (2016).
23. Nothaft, H. & Szymanski, C. M. Protein glycosylation in bacteria: Sweeter than ever. *Nat. Rev. Microbiol.* **8**, 765–778 (2010).
24. Moens, S. & Vanderleyden, J. Glycoproteins in prokaryotes. *Arch. Microbiol.* **168**, 169–175 (1997).
25. Khoury, G. A., Baliban, R. C. & Floudas, C. A. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Sci. Rep.* **1**, 1–5 (2011).
26. Neelamegham, S. *et al.* Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology* **29**, 620–624 (2019).
27. Varki, A. *et al.* Symbol nomenclature for graphical representations of glycans. *Glycobiology* **25**, 1323–1324 (2015).
28. Chen, J. Y. *et al.* Concerted mass spectrometry-based glycomic approach for precision mapping of sulfo sialylated N-glycans on human peripheral blood mononuclear cells and lymphocytes. *Glycobiology* **28**, 9–20 (2018).
29. Coutinho, M. F., Prata, M. J. & Alves, S. Mannose-6-phosphate pathway: A review on its role in lysosomal function and dysfunction. *Mol. Genet. Metab.* **105**, 542–550 (2012).
30. Akella, N. M., Ciraku, L. & Reginato, M. J. Fueling the fire: emerging role of the hexosamine biosynthetic pathway in cancer. *BMC Biol.* 2019 171 **17**, 1–14 (2019).
31. Surleac, M. D. *et al.* The Structural Assessment of Glycosylation Sites Database - SAGS – An Overall View on N-Glycosylation. in *Glycosylation* (ed. Petrescu, S. M.) 3–20 (InTech, 2012). doi:10.5772/51690
32. Nagae, M. & Yamaguchi, Y. Function and 3D Structure of the N-Glycans

- on Glycoproteins. *Int. J. Mol. Sci.* 2012, Vol. 13, Pages 8398-8429 **13**, 8398–8429 (2012).
33. Schulz, B. L. Beyond the Sequon: Sites of N-Glycosylation. in *Glycosylation* (ed. Petrescu, S.) 21–25 (InTech, 2012).
 34. Murray, A. N. *et al.* Enhanced Aromatic Sequons Increase Oligosaccharyltransferase Glycosylation Efficiency and Glycan Homogeneity. *Chem. Biol.* **22**, 1052–1062 (2015).
 35. Stanley, P. N-Glycans. in *Essentials of Glycobiology* (eds. Varki, A. *et al.*) (Cold Spring Harbor Laboratory Press, 2022). doi:10.1016/b978-0-12-821618-7.00047-x
 36. Zielinska, D. F., Gnad, F., Schropp, K., Wiśniewski, J. R. & Mann, M. Mapping N-Glycosylation Sites across Seven Evolutionarily Distant Species Reveals a Divergent Substrate Proteome Despite a Common Core Machinery. *Mol. Cell* **46**, 542–548 (2012).
 37. Croset, A. *et al.* Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. *J. Biotechnol.* **161**, 336–348 (2012).
 38. Butler, M. & Spearman, M. The choice of mammalian cell host and possibilities for glycosylation engineering. *Curr. Opin. Biotechnol.* **30**, 107–112 (2014).
 39. Brodsky, A. N., Caldwell, M. & Harcum, S. W. Glycosylation and post-translational modification gene expression analysis by DNA microarrays for cultured mammalian cells. *Methods* **56**, 408–417 (2012).
 40. Yu, R. K. & Bieberich, E. Regulation of glycosyltransferases in ganglioside biosynthesis by phosphorylation and dephosphorylation. *Mol. Cell. Endocrinol.* **177**, 19–24 (2001).
 41. Mikolajczyk, K., Kaczmarek, R. & Czerwinski, M. How glycosylation affects glycosylation: the role of N-glycans in glycosyltransferase activity. *Glycobiology* **30**, 941–969 (2020).
 42. Agrawal, P. *et al.* Mapping posttranscriptional regulation of the human glycome uncovers microRNA defining the glycode. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4338–4343 (2014).
 43. Guérardel, Y., Chang, L. Y., Maes, E., Huang, C. J. & Khoo, K. H. Glycomic survey mapping of zebrafish identifies unique sialylation pattern. *Glycobiology* **16**, 244–257 (2006).
 44. Varki, A., Freeze, H. H. & Vacquier, V. D. Glycans in Development and Systemic Physiology. in *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, 2009).
 45. Abeln, M. *et al.* Sialylation Is Dispensable for Early Murine Embryonic Development in Vitro. *Chembiochem* **18**, 1305 (2017).
 46. Lefeber, D. J., Freeze, H. H., Steet, R. & Kinoshita, T. Congenital Disorders of Glycosylation. in *Essentials of Glycobiology* 121–125 (Cold Spring Harbor Laboratory Press, 2022). doi:10.1101/GLYCOBIOLOGY.4E.45
 47. Hanson, S. R. *et al.* The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. *Proc. Natl. Acad.*

- Sci. U. S. A.* **106**, 3131–3136 (2009).
48. Mitra, N., Sinha, S., Ramya, T. N. C. & Surolia, A. N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* **31**, 156–163 (2006).
 49. Bosques, C. J. & Imperiali, B. The interplay of glycosylation and disulfide formation influences fibrillization in a prion protein fragment. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 7593–7598 (2003).
 50. Buck, T. M., Eledge, J. & Skach, W. R. Evidence for stabilization of aquaporin-2 folding mutants by N-linked glycosylation in endoplasmic reticulum. *Am. J. Physiol. - Cell Physiol.* **287**, 1292–1299 (2004).
 51. Muthusamy, S. *et al.* N-glycosylation is essential for ileal ASBT function and protection against proteases. *Am. J. Physiol. - Cell Physiol.* **308**, C964–C971 (2015).
 52. Skropeta, D. The effect of individual N-glycans on enzyme activity. *Bioorg. Med. Chem.* **17**, 2645–2653 (2009).
 53. Lai, Y. J. J. *et al.* N-Glycan Branching Affects the Subcellular Distribution of and Inhibition of Matriptase by HAI-2/Placental Bikunin. *PLoS One* **10**, e0132163 (2015).
 54. Han, L. *et al.* The role of N-Glycan modification of TNFR1 in inflammatory microglia activation. *Glycoconj. J.* **32**, 685–693 (2015).
 55. Azimzadeh Irani, M., Kannan, S. & Verma, C. Role of N-glycosylation in EGFR ectodomain ligand binding. *Proteins Struct. Funct. Bioinforma.* **85**, 1529–1549 (2017).
 56. Ioffe, E., Liu, Y. & Stanley, P. Essential role for complex N-glycans in forming an organized layer of bronchial epithelium. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 11041–11046 (1996).
 57. Wang, X. *et al.* Dysregulation of TGF- β 1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15791–15796 (2005).
 58. Rudman, N., Gornik, O. & Lauc, G. Altered N-glycosylation profiles as potential biomarkers and drug targets in diabetes. *FEBS Lett.* **593**, 1598–1615 (2019).
 59. Ohtsubo, K. *et al.* Dietary and Genetic Control of Glucose Transporter 2 Glycosylation Promotes Insulin Secretion in Suppressing Diabetes. *Cell* **123**, 1307–1321 (2005).
 60. Zabczyńska, M., Kozłowska, K. & Pocheć, E. Glycosylation in the Thyroid Gland: Vital Aspects of Glycoprotein Function in Thyrocyte Physiology and Thyroid Disorders. *Int. J. Mol. Sci.* 2018, Vol. 19, Page 2792 **19**, 2792 (2018).
 61. Blazev, R. *et al.* Integrated glycoproteomics identifies a role of N-glycosylation and galectin-1 on myogenesis and muscle development. *Mol. Cell. Proteomics* **20**, 100030 (2021).
 62. Muthusamy, K., Ligezka, A., Johnsen, C., Boyer, S. & Morava-Kozicz, E. Neurological manifestations in PMM2 related congenital disorders of

- glycosylation (CDG): Insights into clinico-radiological characteristics and recommendations for follow-up. *Genet. Med.* **24**, (2022).
63. Ednie, A. R., Parrish, A. R., Sonner, M. J. & Bennett, E. S. Reduced hybrid/complex N-glycosylation disrupts cardiac electrical signaling and calcium handling in a model of dilated cardiomyopathy. *J. Mol. Cell. Cardiol.* **132**, 13–23 (2019).
 64. Clark, M. C. & Baum, L. G. T cells modulate glycans on CD43 and CD45 during development and activation, signal regulation, and survival. *Ann. N. Y. Acad. Sci.* **1253**, 58–67 (2012).
 65. Thiemann, S. & Baum, L. G. Galectins and Immune Responses-Just How Do They Do Those Things They Do? *Annu. Rev. Immunol.* **34**, 243–264 (2016).
 66. Ereño-Orbea, J. *et al.* Molecular basis of human CD22 function and therapeutic targeting. *Nat. Commun.* **2017 81 8**, 1–11 (2017).
 67. O'Keefe, T. L., Williams, G. T., Batista, F. D. & Neuberger, M. S. Deficiency in CD22, a B Cell-specific Inhibitory Receptor, Is Sufficient to Predispose to Development of High Affinity Autoantibodies. *J. Exp. Med.* **189**, 1307–1313 (1999).
 68. Lasky, L. A. Selectin-Carbohydrate Interactions and the Initiation of the Inflammatory Response. *Annu. Rev. Biochem.* **64**, 113–140 (1995).
 69. Stähli, B. E. *et al.* Effects of the P-Selectin Antagonist Inclacumab on Myocardial Damage After Percutaneous Coronary Intervention According to Timing of Infusion: Insights from the SELECT-ACS Trial. *J. Am. Heart Assoc.* **5**, (2016).
 70. Wun, T. *et al.* Phase 1 Study of the E-Selectin Inhibitor GMI 1070 in Patients with Sick Cell Anemia. *PLoS One* **9**, e101301 (2014).
 71. Varki, A. & Kornfeld, S. Historical Background and Overview. in *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, 2017). doi:10.1101/GLYCOBIOLOGY.3E.001
 72. Anderson, N. L. & Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects. *Mol. Cell. Proteomics* **1**, 845–867 (2002).
 73. Qian, W. J. *et al.* Enhanced detection of low abundance human plasma proteins using a tandem IgY12-SuperMix immunoaffinity separation strategy. *Mol. Cell. Proteomics* **7**, 1963–1973 (2008).
 74. Dey, K. K. *et al.* Deep undepleted human serum proteome profiling toward biomarker discovery for Alzheimer's disease. *Clin. Proteomics* **16**, 16 (2019).
 75. Adamczyk, B., Tharmalingam, T. & Rudd, P. M. Glycans as cancer biomarkers. *Biochimica et Biophysica Acta - General Subjects* **1820**, 1347–1353 (2012).
 76. Kailemia, M. J., Park, D. & Lebrilla, C. B. Glycans and glycoproteins as specific biomarkers for cancer. *Anal. Bioanal. Chem.* **409**, 395–410 (2017).
 77. Dotz, V. & Wuhrer, M. N-glycome signatures in human plasma: associations with physiology and major diseases. *FEBS Lett.* **593**, 2966–

- 2976 (2019).
78. Gornik, O. *et al.* Stability of N-glycan profiles in human plasma. *Glycobiology* **19**, 1547–1553 (2009).
 79. Gebrehiwot, A. G. *et al.* Healthy human serum N-glycan profiling reveals the influence of ethnic variation on the identified cancer-relevant glycan biomarkers. *PLoS One* **13**, 1–24 (2018).
 80. Knezevic, A. *et al.* Variability, Heritability and Environmental Determinants of Human Plasma N-Glycome. *J. Proteome Res* **8**, 26 (2009).
 81. Vanhooren, V. *et al.* Serum N-glycan profile shift during human ageing. *Exp. Gerontol.* **45**, 738–743 (2010).
 82. Dall'Olio, F. *et al.* N-glycomic biomarkers of biological aging and longevity: A link with inflammaging. *Ageing Res. Rev.* **12**, 685–698 (2013).
 83. Ruhaak, L. R. *et al.* Plasma protein N-glycan profiles are associated with calendar age, familial longevity and health. *J. Proteome Res.* **10**, 1667–1674 (2011).
 84. Knežević, A. *et al.* Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* **20**, 959–969 (2010).
 85. Ruhaak, L. R., Uh, H. W., Deelder, A. M., Dolhain, R. E. J. M. & Wuhrer, M. Total plasma N-glycome changes during pregnancy. *J. Proteome Res.* **13**, 1657–1668 (2014).
 86. Nilsson, A., Santoro, A., Franceschi, C. & Kadi, F. Detrimental links between physical inactivity, metabolic risk and N-glycomic biomarkers of aging. *Exp. Gerontol.* **124**, 110626 (2019).
 87. Ventham, N. T. *et al.* Changes to Serum Sample Tube and Processing Methodology Does Not Cause Inter-Individual Variation in Automated Whole Serum N-Glycan Profiling in Health and Disease. *PLoS One* **10**, e0123028 (2015).
 88. Stöckmann, H., O'Flaherty, R., Adamczyk, B., Saldova, R. & Rudd, P. M. Automated, high-throughput serum glycoprofiling platform. *Integr. Biol.* **7**, 1026–1032 (2015).
 89. Ruhaak, L. R. *et al.* Optimized workflow for preparation of APTS-labeled N-glycans allowing high-throughput analysis of human plasma glycomes using 48-channel multiplexed CGE-LIF. *J. Proteome Res.* **9**, 6655–6664 (2010).
 90. M Vreeker, G. C. *et al.* Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. (2018). doi:10.1021/acs.analchem.8b02391
 91. Reiding, K. R. *et al.* High-throughput serum N-glycomics: Method comparison and application to study rheumatoid arthritis and pregnancy-associated changes. *Mol. Cell. Proteomics* **18**, 3–15 (2019).
 92. Xie, Y. *et al.* High-throughput and high-sensitivity N-Glycan profiling: A platform for biopharmaceutical development and disease biomarker discovery. *Anal. Biochem.* **623**, 114205 (2021).
 93. Reiding, K. R. *et al.* Human Plasma N-glycosylation as Analyzed by Matrix-

- Assisted Laser Desorption/Ionization-Fourier Transform Ion Cyclotron Resonance-MS Associates with Markers of Inflammation and Metabolic Health. *Mol. Cell. Proteomics* **16**, 228–242 (2017).
94. Bladergroen, M. R. *et al.* Automation of high-throughput mass spectrometry-based plasma n-glycome analysis with linkage-specific sialic acid esterification. *J. Proteome Res.* **14**, 4080–4086 (2015).
 95. Vreeker, G. C. M. *et al.* Dried blood spot N-glycome analysis by MALDI mass spectrometry. *Talanta* **205**, 120104 (2019).
 96. Zhang, Z. *et al.* Serum protein N-glycosylation changes in multiple myeloma. *Biochim. Biophys. Acta - Gen. Subj.* **1863**, 960–970 (2019).
 97. Clerc, F. *et al.* Plasma N-Glycan Signatures Are Associated With Features of Inflammatory Bowel Diseases. *Gastroenterology* **155**, 829–843 (2018).
 98. de Vroome, S. W. *et al.* Serum N-glycome alterations in colorectal cancer associate with survival. *Oncotarget* **9**, 30610–30623 (2018).
 99. Miura, Y. *et al.* BlotGlycoABC™, an integrated glycoblotting technique for rapid and large scale clinical glycomics. *Mol. Cell. Proteomics* **7**, 370–377 (2008).
 100. Nishimura, S. I. Toward automated glycan analysis. *Adv. Carbohydr. Chem. Biochem.* **65**, 219–271 (2011).
 101. Miyahara, K. *et al.* Serum Glycan Markers for Evaluation of Disease Activity and Prediction of Clinical Course in Patients with Ulcerative Colitis. *PLoS One* **8**, e74861 (2013).
 102. Gebrehiwot, A. G. *et al.* Exploring serum and immunoglobulin G N-glycome as diagnostic biomarkers for early detection of breast cancer in Ethiopian women. *BMC Cancer* **19**, 1–18 (2019).
 103. Hatakeyama, S. *et al.* Serum N-Glycan Alteration Associated with Renal Cell Carcinoma Detected by High Throughput Glycan Analysis. *J. Urol.* **191**, 805–813 (2014).
 104. Gizaw, S. T., Ohashi, T., Tanaka, M., Hinou, H. & Nishimura, S. I. Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. *Biochim. Biophys. Acta - Gen. Subj.* **1860**, 1716–1727 (2016).
 105. Matsumoto, T. *et al.* Serum N-glycan profiling is a potential biomarker for castration-resistant prostate cancer. *Sci. Rep.* **9**, 16761 (2019).
 106. Noro, D. *et al.* Serum aberrant N-glycan profile as a marker associated with early antibody-mediated rejection in patients receiving a living donor kidney transplant. *Int. J. Mol. Sci.* **18**, (2017).
 107. Ginsberg, J. M., Chang, B. S., Matarese, R. A. & Garella, S. Use of Single Voided Urine Samples to Estimate Quantitative Proteinuria. *N. Engl. J. Med.* **309**, 1543–1546 (1983).
 108. Ritz, E. & Orth, S. R. Nephropathy in Patients with Type 2 Diabetes Mellitus. *N. Engl. J. Med.* **341**, 1127–1133 (1999).
 109. Currie, G. & Delles, C. Proteinuria and its relation to cardiovascular disease. *Int. J. Nephrol. Renovasc. Dis.* **7**, 13–24 (2013).

110. Soloway, M. S. *et al.* Use of a new tumor marker, urinary NMP22, in the detection of occult or rapidly recurring transitional cell carcinoma of the urinary tract following surgical treatment. *J. Urol.* **156**, 363–367 (1996).
111. Sun, S. *et al.* N-GlycositeAtlas: A database resource for mass spectrometry-based human N-linked glycoprotein and glycosylation site mapping. *Clin. Proteomics* **16**, 1–11 (2019).
112. Li, H., Patel, V., DiMartino, S. E., Froehlich, J. W. & Lee, R. S. An in-depth Comparison of the Pediatric and Adult Urinary N-glycomes. *Mol. Cell. Proteomics* **19**, 1767–1776 (2020).
113. Li, H. *et al.* Uromodulin Isolation and Its N-Glycosylation Analysis by NanoLC-MS/MS. *J. Proteome Res.* **20**, 2662–2672 (2021).
114. Lilja, H., Ulmert, D. & Vickers, A. J. Prostate-specific antigen and prostate cancer: Prediction, detection and monitoring. *Nat. Rev. Cancer* **8**, 268–278 (2008).
115. Barry, M. J. Screening for Prostate Cancer — The Controversy That Refuses to Die. *N. Engl. J. Med.* **360**, 1351–1354 (2009).
116. Kim, Y. *et al.* Identification of differentially expressed proteins in direct expressed prostatic secretions of men with organ-confined versus extracapsular prostate cancer. *Mol. Cell. Proteomics* **11**, 1870–1884 (2012).
117. Tkac, J. *et al.* Glycomics of prostate cancer: updates. *Expert Review of Proteomics* **16**, 65–76 (2019).
118. Llop, E. *et al.* Improvement of Prostate Cancer Diagnosis by Detecting PSA Glycosylation-Specific Changes. *Theranostics* **6**, 1190–1204 (2016).
119. Gilgunn, S. *et al.* Glycosylation in indolent, significant and aggressive prostate cancer by automated high-throughput n-glycan profiling. *Int. J. Mol. Sci.* **21**, 1–17 (2020).
120. Murphy, K. *et al.* Integrating biomarkers across omic platforms: an approach to improve stratification of patients with indolent and aggressive prostate cancer. *Mol. Oncol.* **12**, 1513–1525 (2018).
121. Ishibashi, Y. *et al.* Serum tri- and tetra-antennary N-glycan is a potential predictive biomarker for castration-resistant prostate cancer. *Prostate* **74**, 1521–1529 (2014).
122. Saldova, R., Fan, Y., Fitzpatrick, J. M., Watson, R. W. G. & Rudd, P. M. Core fucosylation and α 2-3 sialylation in serum N-glycome is significantly increased in prostate cancer comparing to benign prostate hyperplasia. *Glycobiology* **21**, 195–205 (2011).
123. Vermassen, T. *et al.* Urinary prostate protein glycosylation profiling as a diagnostic biomarker for prostate cancer. *Prostate* **75**, 314–322 (2015).
124. Vermassen, T. *et al.* Capillary electrophoresis of urinary prostate glycoproteins assists in the diagnosis of prostate cancer. *Electrophoresis* **35**, 1017–1024 (2014).
125. Nyalwidhe, J. O. *et al.* Increased bisecting N-acetylglucosamine and decreased branched chain glycans of N-linked glycoproteins in expressed prostatic secretions associated with prostate cancer progression.

- Proteomics - Clin. Appl.* **7**, 677–689 (2013).
126. Litman, G. W., Anderson, M. K. & Rast, J. P. Evolution of antigen binding receptors. *Annu. Rev. Immunol.* **17**, 109–147 (1999).
 127. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M. & Dwek, R. A. The Impact of Glycosylation on the Biological Function and Structure of Human Immunoglobulins. *Annu. Rev. Immunol.* **25**, 21–50 (2007).
 128. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: From structure to effector functions. *Front. Immunol.* **5**, 520 (2014).
 129. Russell, A., Adua, E., Ugrina, I., Laws, S. & Wang, W. Unravelling immunoglobulin G Fc N-glycosylation: A dynamic marker potentiating predictive, preventive and personalised medicine. *Int. J. Mol. Sci.* **19**, 390 (2018).
 130. Anumula, K. R. Quantitative glycan profiling of normal human plasma derived immunoglobulin and its fragments Fab and Fc. *J. Immunol. Methods* **382**, 167–176 (2012).
 131. Zhu, D. *et al.* Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. *Blood* **99**, 2562–2568 (2002).
 132. Dunn-Walters, D., Boursier, L. & Spencer, J. Effect of somatic hypermutation on potential N-glycosylation sites in human immunoglobulin heavy chain variable regions. *Mol. Immunol.* **37**, 107–113 (2000).
 133. Gudelj, I., Lauc, G. & Pezer, M. Immunoglobulin G glycosylation in aging and diseases. *Cell. Immunol.* **333**, 65–79 (2018).
 134. van de Bovenkamp, F. S., Hafkenscheid, L., Rispens, T. & Rombouts, Y. The Emerging Importance of IgG Fab Glycosylation in Immunity. *J. Immunol.* **196**, 1435–1441 (2016).
 135. Samuelsson, A., Towers, T. L. & Ravetch, J. V. Anti-inflammatory activity of IVIG mediated through the inhibitory Fc receptor. *Science (80-.)*. **291**, 484–486 (2001).
 136. Kanda, Y. *et al.* Comparison of biological activity among nonfucosylated therapeutic IgG1 antibodies with three different N-linked Fc oligosaccharides: The high-mannose, hybrid, and complex types. *Glycobiology* **17**, 104–118 (2007).
 137. Masuda, K. *et al.* Enhanced binding affinity for FcγRIIIa of fucose-negative antibody is sufficient to induce maximal antibody-dependent cellular cytotoxicity. *Mol. Immunol.* **44**, 3122–3131 (2007).
 138. Banda, N. K. *et al.* Initiation of the alternative pathway of murine complement by immune complexes is dependent on N-glycans in IgG antibodies. *Arthritis Rheum.* **58**, 3081–3089 (2008).
 139. Malhotra, R. *et al.* Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. *Nat. Med.* **1**, 237–243 (1995).
 140. Arnold, J. N., Dwek, R. A., Rudd, P. M. & Sim, R. B. Mannan binding lectin and its interaction with immunoglobulins in health and in disease. *Immunol. Lett.* **106**, 103–110 (2006).

141. Nimmerjahn, F., Anthony, R. M. & Ravetch, J. V. Agalactosylated IgG antibodies depend on cellular Fc receptors for in vivo activity. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8433–8437 (2007).
142. Krištić, J., Lauc, G. & Pezer, M. Immunoglobulin G glycans – Biomarkers and molecular effectors of aging. *Clin. Chim. Acta* **535**, 30–45 (2022).
143. Huffman, J. E. *et al.* Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research. *Mol. Cell. Proteomics* **13**, 1598–1610 (2014).
144. Falck, D., Jansen, B. C., de Haan, N. & Wuhrer, M. High-throughput analysis of IgG Fc glycopeptides by LC-MS. *Methods Mol. Biol.* **1503**, 31–47 (2017).
145. Stark, G. B., Grandel, S. & Spilker, G. Tissue suction of the male and female breast. *Aesthetic Plast. Surg.* 1992 164 **16**, 317–324 (1992).
146. Akram, M., Iqbal, M., Daniyal, M. & Khan, A. U. Awareness and current knowledge of breast cancer. *Biol. Res.* 2017 501 **50**, 1–23 (2017).
147. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **71**, 209–249 (2021).
148. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA. Cancer J. Clin.* **72**, 7–33 (2022).
149. Arnold, M. *et al.* Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast Off. J. Eur. Soc. Mastology* **66**, 15 (2022).
150. Harbeck, N. *et al.* Breast Cancer. *Nat. Rev. Dis. Prim.* **66**, (2019).
151. Engel, C. & Fischer, C. Breast Cancer Risks and Risk Prediction Models. *Breast Care* **10**, 7–12 (2015).
152. Paluch-Shimon, S. *et al.* Prevention and screening in BRCA mutation carriers and other breast/ovarian hereditary cancer syndromes: ESMO clinical practice guidelines for cancer prevention and screening. *Ann. Oncol.* **27**, v103–v110 (2016).
153. Shiovitz, S. & Korde, L. A. Genetics of breast cancer: a topic in evolution. *Ann. Oncol.* **26**, 1291 (2015).
154. Brewer, H. R., Jones, M. E., Schoemaker, M. J., Ashworth, A. & Swerdlow, A. J. Family history and risk of breast cancer: an analysis accounting for family structure. *Breast Cancer Res. Treat.* **165**, 193–200 (2017).
155. DeSantis, C. E. *et al.* Breast cancer statistics, 2019. *CA. Cancer J. Clin.* **69**, 438–451 (2019).
156. Horn, J. & Vatten, L. J. Reproductive and hormonal risk factors of breast cancer: A historical perspective. *Int. J. Womens. Health* **9**, 265–272 (2017).
157. Horn, J., Åsvold, B. O., Opdahl, S., Tretli, S. & Vatten, L. J. Reproductive factors and the risk of breast cancer in old age: A Norwegian cohort study. *Breast Cancer Res. Treat.* **139**, 237–243 (2013).
158. Rosato, V. *et al.* Reproductive and hormonal factors, family history, and breast cancer according to the hormonal receptor status. *Eur. J. Cancer Prev.* **23**, 412–417 (2014).

159. Del Pup, L., Codacci-Pisanelli, G. & Peccatori, F. Breast cancer risk of hormonal contraception: Counselling considering new evidence. *Crit. Rev. Oncol. Hematol.* **137**, 123–130 (2019).
160. Mørch, L. S. *et al.* Contemporary Hormonal Contraception and the Risk of Breast Cancer. *N. Engl. J. Med.* **377**, 2228–2239 (2017).
161. Busund, M. *et al.* Progestin-only and combined oral contraceptives and receptor-defined premenopausal breast cancer risk: The Norwegian Women and Cancer Study. *Int. J. Cancer* **142**, 2293–2302 (2018).
162. Advani, S. M. *et al.* Association of Breast Density With Breast Cancer Risk Among Women Aged 65 Years or Older by Age Group and Body Mass Index. *JAMA Netw. Open* **4**, e2122810–e2122810 (2021).
163. Seiler, A., Chen, M. A., Brown, R. L. & Fagundes, C. P. Obesity, Dietary Factors, Nutrition, and Breast Cancer Risk. *Curr. Breast Cancer Reports 2018 101* **10**, 14–27 (2018).
164. Gaudet, M. M. *et al.* Pooled analysis of nine cohorts reveals breast cancer risk factors by tumor molecular subtype. *Cancer Res.* **78**, 6011–6021 (2018).
165. Nielsen, T. O. *et al.* Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *J. Natl. Cancer Inst.* **113**, 808–819 (2021).
166. Koboldt, D. C. *et al.* Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
167. Cheang, M. C. U. *et al.* Defining Breast Cancer Intrinsic Subtypes by Quantitative Receptor Expression. *Oncologist* **20**, 474–482 (2015).
168. Spak, D. A., Plaxco, J. S., Santiago, L., Dryden, M. J. & Dogan, B. E. BI-RADS® fifth edition: A summary of changes. *Diagn. Interv. Imaging* **98**, 179–190 (2017).
169. Pace, L. E. & Keating, N. L. A systematic assessment of benefits and risks to guide breast cancer screening decisions. *JAMA - Journal of the American Medical Association* **311**, 1327–1335 (2014).
170. Welch, H. G. & Black, W. C. Overdiagnosis in cancer. *J. Natl. Cancer Inst.* **102**, 605–613 (2010).
171. Lehman, C. D. *et al.* National performance benchmarks for modern screening digital mammography: Update from the Breast Cancer Surveillance Consortium. *Radiology* **283**, 49–58 (2017).
172. US Food and Drug Administration. MQSA National Statistics. *MQSA Insights* (2022).
173. Schopper, D. & de Wolf, C. How effective are breast cancer screening programmes by mammography? Review of the current evidence. *Eur. J. Cancer* **45**, 1916–1923 (2009).
174. Tosteson, A. N. A. *et al.* Consequences of False-Positive Screening Mammograms. *JAMA Intern. Med.* **174**, 954–961 (2014).
175. Haas, J. S. *et al.* Disparities in the use of screening magnetic resonance imaging of the breast in community practice by race, ethnicity, and socioeconomic status. *Cancer* **122**, 611–617 (2016).

176. Lee, J. M. *et al.* Performance of Screening Ultrasonography as an Adjunct to Screening Mammography in Women Across the Spectrum of Breast Cancer Risk. *JAMA Intern. Med.* **179**, 658–667 (2019).
177. Howlader, N. *et al.* *SEER Cancer Statistics Review, 1975-2017 SEER Cancer Statistics.* National Cancer Institute (2017).
178. Duffy, M. J., Shering, S., Sherry, F., McDermott, E. & O'Higgins, N. CA 15-3: A prognostic marker in breast cancer. *Int. J. Biol. Markers* **15**, 330–333 (2000).
179. Duffy, M. J. Serum Tumor Markers in Breast Cancer: Are They of Clinical Value? *Clin. Chem.* **52**, 345–351 (2006).
180. Li, X. *et al.* Clinicopathological and prognostic significance of cancer antigen 15-3 and carcinoembryonic antigen in breast cancer: A Meta-Analysis including 12,993 Patients. *Dis. Markers* **2018**, 1–15 (2018).
181. Ciborowski, P. & Finn, O. J. Non-glycosylated tandem repeats of MUC1 facilitate attachment of breast tumor cells to normal human lung tissue and immobilized extracellular matrix proteins (ECM) in vitro: Potential role in metastasis. *Clin. Exp. Metastasis* **19**, 339–345 (2002).
182. Ghosh, S. K., Pantazopoulos, P., Medarova, Z. & Moore, A. Expression of underglycosylated MUC1 antigen in cancerous and adjacent normal breast tissues. *Clin. Breast Cancer* **13**, 109–118 (2013).
183. Klee, G. G. & Schreiber, W. E. MUC1 gene-derived glycoprotein assays for monitoring breast cancer (CA 15-3, CA 27.29, BR): Are they measuring the same antigen? *Arch. Pathol. Lab. Med.* **128**, 1131–1135 (2004).
184. Hammarström, S. The carcinoembryonic antigen (CEA) family: Structures, suggested functions and expression in normal and malignant tissues. *Semin. Cancer Biol.* **9**, 67–81 (1999).
185. Ali, S. M. *et al.* Serum HER-2/neu and relative resistance to trastuzumab-based therapy in patients with metastatic breast cancer. *Cancer* **113**, 1294–1301 (2008).
186. Perrier, A. *et al.* An updated evaluation of serum sHER2, CA15.3, and CEA levels as biomarkers for the response of patients with metastatic breast cancer to trastuzumab-based therapies. *PLoS One* **15**, e0227356 (2020).
187. Christianson, T. A. *et al.* NH₂-terminally truncated HER-2/neu protein: Relationship with shedding of the extracellular domain and with prognostic factors in breast cancer. *Cancer Res.* **58**, 5123–5129 (1998).
188. Bon, G. G. *et al.* *Clinical and technical evaluation of ACS BR serum assay of MUC1 gene-derived glycoprotein in breast cancer, and comparison with CA 15-3 assays.*
189. Stieber, P. *et al.* Clinical evaluation of the Elecsys CA 15-3 test in breast cancer patients. *Clin. Lab.* **49**, 15–24 (2003).
190. Hayes, D. F., Zurawski, V. R. & Kufe, D. W. Comparison of circulating CA15-3 and carcinoembryonic antigen levels in patients with breast cancer. *J. Clin. Oncol.* **4**, 1542–1550 (1986).
191. Swellam, M. *et al.* Aberrant methylation of APC and RAR β 2 genes in breast cancer patients. *IUBMB Life* **67**, 61–68 (2015).

192. Gao, J., Zhang, Q., Xu, J., Guo, L. & Li, X. Clinical significance of serum miR-21 in breast cancer compared with CA153 and CEA. *Chinese J. Cancer Res.* **25**, 743–748 (2013).
193. Streckfus, C. *et al.* The presence of soluble c-erbB-2 in saliva and serum among women with breast carcinoma: A preliminary study. *Clin. Cancer Res.* **6**, 2363–2370 (2000).
194. Ashkani, J. & Naidoo, K. J. Glycosyltransferase Gene Expression Profiles Classify Cancer Types and Propose Prognostic Subtypes. *Sci. Rep.* **6**, 1–8 (2016).
195. Milde-Langosch, K. *et al.* Prognostic relevance of glycosylation-associated genes in breast cancer. *Breast Cancer Res. Treat.* **145**, 295–305 (2014).
196. Legler, K. *et al.* Reduced mannosidase MAN1A1 expression leads to aberrant N-glycosylation and impaired survival in breast cancer. *Br. J. Cancer* **2018 1186** **118**, 847–856 (2018).
197. Peng, W., zhu, R., Zhou, S., Mirzaei, P. & Mechref, Y. Integrated Transcriptomics, Proteomics, and Glycomics Reveals the Association between Up-regulation of Sialylated N-glycans/Integrin and Breast Cancer Brain Metastasis. *Sci. Rep.* **9**, 12–14 (2019).
198. Lin, S., Kemmner, W., Grigull, S. & Schlag, P. M. Cell surface $\alpha 2$, 6-sialylation affects adhesion of breast carcinoma cells. *Exp. Cell Res.* **276**, 101–110 (2002).
199. Cui, H., Lin, Y., Yue, L., Zhao, X. & Liu, J. Differential expression of the $\alpha 2,3$ -sialic acid residues in breast cancer is associated with metastatic potential. *Oncol. Rep.* **25**, 1365–1371 (2011).
200. Ščupáková, K. *et al.* Clinical importance of high-mannose, fucosylated, and complex N-glycans in breast cancer metastasis. *JCI Insight* **6**, (2021).
201. Li, Q. *et al.* Comprehensive N-Glycome Profiling of Cells and Tissues for Breast Cancer Diagnosis. *J. Proteome Res.* **18**, 2559–2570 (2019).
202. Yaman, M. E., Kayili, H. M., Albayrak, M., Kadioglu, Y. & Salih, B. Differential N-glycosylation profiling of formalin-fixed paraffin-embedded (FFPE) invasive ductal carcinoma tissues using MALDI-TOF-MS. *Mol. Omi.* **17**, 394–404 (2021).
203. Scott, D. A. *et al.* Increases in Tumor N-Glycan Polylactosamines Associated with Advanced HER2-Positive and Triple-Negative Breast Cancer Tissues. *PROTEOMICS – Clin. Appl.* **13**, 1800014 (2019).
204. Yang, Z., Harris, L. E., Palmer-Toy, D. E. & Hancock, W. S. Multilectin Affinity Chromatography for Characterization of Multiple Glycoprotein Biomarker Candidates in Serum from Breast Cancer Patients. *Clin. Chem.* **52**, 1897–1905 (2006).
205. Kirmiz, C. *et al.* A Serum Glycomics Approach to Breast Cancer Biomarkers. (2007). doi:10.1074/mcp.M600171-MCP200
206. Kyselova, Z. *et al.* Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin. Chem.* **54**, 1166–1175 (2008).

207. Abd Hamid, U. M. *et al.* A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression. *Glycobiology* **18**, 1105–1118 (2008).
208. Alley, W. R., Madera, M., Mechref, Y. & Novotny, M. V. Chip-based reversed-phase liquid chromatography-mass spectrometry of permethylated N-linked glycans: A potential methodology for cancer-biomarker discovery. *Anal. Chem.* **82**, 5095–5106 (2010).
209. Alley, W. R. & Novotny, M. V. Glycomic analysis of sialic acid linkages in glycans derived from blood serum glycoproteins. *J. Proteome Res.* **9**, 3062–3072 (2010).
210. de Leoz, M. L. A. *et al.* High-Mannose Glycans are Elevated during Breast Cancer Progression. *Mol. Cell. Proteomics* **10**, 1–9 (2011).
211. Zheng, T., Peelen, D. & Smith, L. M. Lectin arrays for profiling cell surface carbohydrate expression. *J. Am. Chem. Soc.* **127**, 9982–9983 (2005).
212. Pilobello, K. T., Krishnamoorthy, L., Slawek, D. & Mahal, L. K. Development of a Lectin Microarray for the Rapid Analysis of Protein Glycopatterns. doi:10.1002/cbic.200400403
213. Kuno, A. *et al.* Evanescent-field fluorescence-assisted lectin microarray: A new strategy for glycan profiling. *Nat. Methods* **2**, 851–856 (2005).
214. Angeloni, S. *et al.* Glycoprofiling with micro-arrays of glycoconjugates and lectins. *Glycobiology* **15**, 31–41 (2005).
215. Fry, S. A. *et al.* Lectin microarray profiling of metastatic breast cancers. (2011). doi:10.1093/glycob/cwr045
216. Pierce, A. *et al.* Levels of specific glycans significantly distinguish lymph node-positive from lymph node-negative breast cancer patients. *Glycobiology* **20**, 1283–1288 (2010).
217. Saldova, R., Reuben, J. M., Abd Hamid, U. M., Rudd, P. M. & Cristofanilli, M. Levels of specific serum N-glycans identify breast cancer patients with higher circulating tumor cell counts. *Ann. Oncol.* **22**, 1113–1119 (2011).
218. Saldova, R. *et al.* Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC. *J. Proteome Res.* **13**, 2314–2327 (2014).
219. Ju, L. *et al.* Elevated level of serum glycoprotein bifucosylation and prognostic value in Chinese breast cancer. *Glycobiology* **26**, 460–471 (2016).
220. Lee, S. B. *et al.* Breast cancer diagnosis by analysis of serum N-glycans using MALDI-TOF mass spectroscopy. *PLoS One* **15**, 1–11 (2020).
221. D., K., Behan, J., Matthews-Smith, G. & M., A. Alpha-1-Acid Glycoprotein (AGP) as a Potential Biomarker for Breast Cancer. in *Glycosylation* (InTech, 2012). doi:10.5772/48177
222. Carlsson, M. C. *et al.* Galectin-1-Binding Glycoforms of Haptoglobin with Altered Intracellular Trafficking, and Increase in Metastatic Breast Cancer Patients. *PLoS One* **6**, e26560 (2011).
223. Dargan, E., Thompson, S., Cantwell, B. M. J., Wilson, R. G. & Turner, G. A. Changes in the fucose content of haptoglobin in breast and ovarian

- cancer: Association with disease progression. *Glycosylation Dis.* **1**, 37–43 (1994).
224. Zeng, Z. *et al.* The development of an integrated platform to identify breast cancer glycoproteome changes in human serum. *J. Chromatogr. A* **1217**, 3307–3315 (2010).
225. Goodarzi, M. T. & Turner, G. A. Decreased branching, increased fucosylation and changed sialylation of alpha-1-proteinase inhibitor in breast and ovarian cancer. *Clin. Chim. Acta* **236**, 161–171 (1995).
226. Thompson, D. K., Haddow, J. E., Smith, D. E. & Ritchie, R. F. Elevated serum acute phase protein levels as predictors of disseminated breast cancer. *Cancer* **51**, 2100–2104 (1983).
227. Kawaguchi-Sakita, N. *et al.* Serum immunoglobulin G Fc region N-glycosylation profiling by matrix-assisted laser desorption/ionization mass spectrometry can distinguish breast cancer patients from cancer-free controls. *Biochem. Biophys. Res. Commun.* **469**, 1140–1145 (2016).
228. Haakensen, V. D. *et al.* Serum N-glycan analysis in breast cancer patients - Relation to tumour biology and clinical outcome. *Mol. Oncol.* **10**, 59–72
229. Vreeker, G. C. M. *et al.* Serum N-glycan profiles differ for various breast cancer subtypes. *Glycoconj. J.* **38**, 387–395 (2021).
230. Pace, L. E. & Keating, N. L. A Systematic Assessment of Benefits and Risks to Guide Breast Cancer Screening Decisions. *JAMA* **311**, 1327–1335 (2014).
231. Doherty, M. *et al.* Plasma N-glycans in colorectal cancer risk. *Sci. Rep.* **8**, 1–12 (2018).
232. Albrecht, S., Unwin, L., Muniyappa, M. & Rudd, P. M. Glycosylation as a marker for inflammatory arthritis. *Cancer Biomarkers* **14**, 17–28 (2014).
233. Coss, K. P. *et al.* N-glycan abnormalities in children with galactosemia. *J. Proteome Res.* **13**, 385–394 (2014).
234. Tudor, L. *et al.* N-glycomic profile in combat related post-traumatic stress disorder. *Biomolecules* **9**, (2019).
235. Park, D. I. *et al.* Blood plasma/IgG N-glycome biosignatures associated with major depressive disorder symptom severity and the antidepressant response. *Sci. Rep.* **8**, (2018).
236. Hanić, M., Lauc, G. & Trbojević-Akmačić, I. N-Glycan Analysis by Ultra-Performance Liquid Chromatography and Capillary Gel Electrophoresis with Fluorescent Labeling. *Curr. Protoc. Protein Sci.* **97**, (2019).
237. Adamczyk, B., Stöckmann, H., O’Flaherty, R., Karlsson, N. G. & Rudd, P. M. High-throughput analysis of the plasma N-glycome by UHPLC. in *Methods in Molecular Biology* **1503**, 97–108 (Humana Press Inc., 2017).
238. Ruhaak, L. R., Xu, G., Li, Q., Goonatileke, E. & Lebrilla, C. B. Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem. Rev.* **118**, 7886–7930 (2018).
239. Suttapitugsakul, S., Sun, F. & Wu, R. Recent Advances in Glycoproteomic Analysis by Mass Spectrometry. *Anal. Chem.* **92**, 267–291 (2019).

240. Apweiler, R., Hermjakob, H. & Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta - Gen. Subj.* **1473**, 4–8 (1999).
241. Lauc, G., Pezer, M., Rudan, I. & Campbell, H. Mechanisms of disease: The human N-glycome. *Biochim. Biophys. Acta - Gen. Subj.* **1860**, 1574–1582 (2016).
242. Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: Diversity, synthesis and function. *Nature Reviews Molecular Cell Biology* **13**, 448–462 (2012).
243. Tanaka, T. *et al.* Aberrant N-glycosylation profile of serum immunoglobulins is a diagnostic biomarker of urothelial carcinomas. *Int. J. Mol. Sci.* **18**, 1–14 (2017).
244. Powers, T. W. *et al.* MALDI imaging mass spectrometry profiling of N-glycans in formalin-fixed paraffin embedded clinical tissue blocks and tissue microarrays. *PLoS One* **9**, 1–11 (2014).
245. Drake, R. R. *et al.* MALDI Mass Spectrometry Imaging of N-Linked Glycans in Cancer Tissues. in *Advances in Cancer Research* **134**, 85–116 (Academic Press Inc., 2017).
246. Powers, T. W. *et al.* Matrix assisted laser desorption ionization imaging mass spectrometry workflow for spatial profiling analysis of N-linked Glycan expression in tissues. *Anal. Chem.* **85**, 9799–9806 (2013).
247. Drake, R. R., Powers, T. W., Norris-Caneda, K., Mehta, A. S. & Angel, P. M. In Situ Imaging of N-Glycans by MALDI Imaging Mass Spectrometry of Fresh or Formalin-Fixed Paraffin-Embedded Tissue. *Curr. Protoc. Protein Sci.* **94**, 1–21 (2018).
248. Black, A. P. *et al.* A novel mass spectrometry platform for multiplexed N-glycoprotein biomarker discovery from patient biofluids by antibody panel based N-glycan imaging. *Anal. Chem.* **91**, 8429–8435 (2019).
249. West, C. A., Liang, H., Drake, R. R. & Mehta, A. S. A New Enzymatic Approach to Distinguish Fucosylation Isomers of N-linked Glycans in Tissues Using MALDI Imaging Mass Spectrometry. *J. Proteome Res.* (2020). doi:10.1021/acs.jproteome.0c00024
250. Semmes, O. J. *et al.* Evaluation of Serum Protein Profiling by Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry for the Detection of Prostate Cancer: I. Assessment of Platform Reproducibility. *Clin. Chem.* **51**, 102–112 (2005).
251. Schaub, N. P. *et al.* Serum Proteomic Biomarker Discovery Reflective of Stage and Obesity in Breast Cancer Patients. *J. Am. Coll. Surg.* **208**, 970–978 (2009).
252. Saldova, R. *et al.* Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. *Glycobiology* **17**, 1344–1356 (2007).
253. Reiding, K. R., Blank, D., Kuijper, D. M., Deelder, A. A. & Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. (2014).

doi:10.1021/ac500335t

254. Drake, R. R. & Kislinger, T. The proteomics of prostate cancer exosomes. *Expert Review of Proteomics* **11**, 167–177 (2014).
255. Nawaz, M. *et al.* The emerging role of extracellular vesicles as biomarkers for urogenital cancers. *Nature Reviews Urology* **11**, 688–701 (2014).
256. Drake, R. R. *et al.* Clinical collection and protein properties of expressed prostatic secretions as a source for biomarkers of prostatic disease. *J. Proteomics* **72**, 907–917 (2009).
257. Drake, R. R. *et al.* In-depth proteomic analyses of direct expressed prostatic secretions. *J. Proteome Res.* **9**, 2109–2116 (2010).
258. Principe, S. *et al.* Identification of prostate-enriched proteins by in-depth proteomic analyses of expressed prostatic secretions in urine. *J. Proteome Res.* **11**, 2386–2396 (2012).
259. Kim, Y. *et al.* Targeted proteomics identifies liquid-biopsy signatures for extracapsular prostate cancer. *Nat. Commun.* **7**, 1–10 (2016).
260. Otto, J. J. *et al.* Targeted Mass Spectrometry of a Clinically Relevant PSA Variant from Post-DRE Urines for Quantitation and Genotype Determination. *Proteomics - Clin. Appl.* **14**, 2000012 (2020).
261. Van Gils, M. P. M. Q. *et al.* Molecular PCA3 diagnostics on prostatic fluid. *Prostate* **67**, 881–887 (2007).
262. Laxman, B. *et al.* A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. *Cancer Res.* **68**, 645–649 (2008).
263. Linxweiler, J. & Junker, K. Extracellular vesicles in urological malignancies: an update. *Nature Reviews Urology* **17**, 11–27 (2020).
264. Erdbrügger, U. *et al.* Urinary extracellular vesicles: A position paper by the Urine Task Force of the International Society for Extracellular Vesicles. *J. Extracell. Vesicles* **10**, e12093 (2021).
265. Wang, Y. T. *et al.* Proteomic analysis of exosomes for discovery of protein biomarkers for prostate and bladder cancer. *Cancers* **12**, 1–19 (2020).
266. Drake, R. R., Jones, E. E., Powers, T. W. & Nyalwidhe, J. O. Altered glycosylation in prostate cancer. in *Advances in Cancer Research* **126**, 345–382 (Academic Press Inc., 2015).
267. Scott, E. & Munkley, J. Glycans as biomarkers in prostate cancer. *Int. J. Mol. Sci.* **20**, 1389 (2019).
268. Saraswat, M. *et al.* N-linked (N-) glycoproteomics of urinary exosomes. *Molecular and Cellular Proteomics* **14**, 2298 (2015).
269. Brown, C. J. *et al.* Glycoproteomic Analysis of Human Urinary Exosomes. *Anal. Chem.* **92**, 14357–14365 (2020).
270. Leymarie, N. *et al.* Interlaboratory study on differential analysis of protein glycosylation by mass spectrometry: The ABRF glycoprotein research multi-institutional study 2012. *Mol. Cell. Proteomics* **12**, 2935–2951 (2013).
271. White, K. Y. *et al.* Glycomic characterization of prostate-specific antigen and prostatic acid phosphatase in prostate cancer and benign disease seminal plasma fluids. *J. Proteome Res.* **8**, 620–630 (2009).
272. Hatakeyama, S., Yoneyama, T., Tobisawa, Y., Yamamoto, H. & Ohyama,

- C. Narrative review of urinary glycan biomarkers in prostate cancer. *Translational Andrology and Urology* **10**, 1850 (2021).
273. Jia, G. *et al.* Alterations in expressed prostate secretion-urine PSA N-glycosylation discriminate prostate cancer from benign prostate hyperplasia. *Oncotarget* **8**, 76987–76999 (2017).
274. Kammeijer, G. S. M., Nouta, J., De La Rosette, J. J. M. C. H., De Reijke, T. M. & Wuhrer, M. An In-Depth Glycosylation Assay for Urinary Prostate-Specific Antigen. *Anal. Chem.* **90**, 4414–4421 (2018).
275. Hanzawa, K. *et al.* Investigation of acidic free-glycans in urine and their alteration in cancer. *Glycobiology* **31**, 391–409 (2021).
276. Song, W. *et al.* In-Depth Compositional and Structural Characterization of N-Glycans Derived from Human Urinary Exosomes. *Anal. Chem.* **91**, 13528–13537 (2019).
277. Kawakita, C. *et al.* Novel urinary glycan profiling by lectin array serves as the biomarkers for predicting renal prognosis in patients with IgA nephropathy. *Sci. Reports 2021 111* **11**, 1–11 (2021).
278. Mise, K. *et al.* Novel Urinary Glycan Biomarkers Predict Cardiovascular Events in Patients With Type 2 Diabetes: A Multicenter Prospective Study With 5-Year Follow Up (U-CARE Study 2). *Front. Cardiovasc. Med.* **0**, 373 (2021).
279. Blaschke, C. R. K., Black, A. P., Mehta, A. S., Angel, P. M. & Drake, R. R. Rapid N-Glycan Profiling of Serum and Plasma by a Novel Slide-Based Imaging Mass Spectrometry Workflow. *J. Am. Soc. Mass Spectrom.* **31**, 2511–2520 (2020).
280. McDowell, C. T. *et al.* Imaging mass spectrometry and lectin analysis of n-linked glycans in carbohydrate antigen-defined pancreatic cancer tissues. *Mol. Cell. Proteomics* **20**, 100012 (2021).
281. Drake, R. R., West, C. A., Mehta, A. S. & Angel, P. M. MALDI mass spectrometry imaging of N-linked glycans in tissues. in *Advances in Experimental Medicine and Biology* **1104**, 59–76 (Springer New York LLC, 2018).
282. Zou, G., Benktander, J. D., Gizaw, S. T., Gaunitz, S. & Novotny, M. V. Comprehensive Analytical Approach toward Glycomic Characterization and Profiling in Urinary Exosomes. *Anal. Chem.* **89**, 5364–5372 (2017).
283. Nelson, H. D. *et al.* Effectiveness of breast cancer screening: Systematic review and meta-analysis to update the 2009 U.S. Preventive services task force recommendation. *Ann. Intern. Med.* **164**, 244–255 (2016).
284. Ong, M. S. & Mandl, K. D. National expenditure for false-positive mammograms and breast cancer overdiagnoses estimated at \$ 4 billion a year. *Health Aff.* **34**, 576–583 (2015).
285. Vlahiotis, A., Griffin, B., Stavros, A. T. & Margolis, J. Analysis of utilization patterns and associated costs of the breast imaging and diagnostic procedures after screening mammography. *Clin. Outcomes Res.* **10**, 157 (2018).
286. Chubak, J., Boudreau, D. M., Fishman, P. A. & Elmore, J. G. Cost of

- Breast-Related Care in the Year Following False Positive Screening Mammograms. *Med. Care* **48**, 815 (2010).
287. Kolb, T. M., Lichy, J. & Newhouse, J. H. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: An analysis of 27,825 patient evaluations. *Radiology* **225**, 165–175 (2002).
 288. Boyd, N. F. *et al.* Mammographic density: The risk and detection of breast cancer. *J. BioLaw Bus.* **10**, 60–61 (2007).
 289. Sechopoulos, I. A review of breast tomosynthesis. Part I. The image acquisition process. *Med. Phys.* **40**, 014301 (2013).
 290. Olgar, T., Kahn, T. & Gosch, D. Average glandular dose in digital mammography and breast tomosynthesis. *RoFo* **184**, 911–918 (2012).
 291. Berg, W. A. *et al.* Reasons Women at Elevated Risk of Breast Cancer Refuse Breast MR Imaging Screening: ACRIN 6666. *Radiology* **254**, 79 (2010).
 292. Evans, A. *et al.* Breast ultrasound: recommendations for information to women and referring physicians by the European Society of Breast Imaging. *Insights Imaging* **9**, 449–461 (2018).
 293. Higel, F. *et al.* N-glycans of complex glycosylated biopharmaceuticals and their impact on protein clearance. *Eur. J. Pharm. Biopharm.* **139**, 123–131 (2019).
 294. Bas, M. *et al.* Fc Sialylation Prolongs Serum Half-Life of Therapeutic Antibodies. *J. Immunol.* **202**, 1582–1594 (2019).
 295. Baynes, J. W. & Wold, F. Effect of glycosylation on the in vivo circulating half-life of ribonuclease. *J. Biol. Chem.* **251**, 6016–6024 (1976).
 296. Blaschke, C. R. K. *et al.* Direct N-Glycosylation Profiling of Urine and Prostatic Fluid Glycoproteins and Extracellular Vesicles. *Front. Chem.* **9**, (2021).
 297. Blaschke, C. R. K. & Drake, R. R. MALDI Mass Spectrometry Imaging and Glycomics. in *MALDI Mass Spectrometry Imaging: From Fundamentals to Spatial Omics* (ed. Porta Siegel, T.) 207–233 (The Royal Society of Chemistry, 2021).
 298. Siu, A. L. Screening for breast cancer: U.S. Preventive services task force recommendation statement. *Ann. Intern. Med.* **164**, 279–296 (2016).
 299. Song, T., Aldredge, D. & Lebrilla, C. B. A Method for In-Depth Structural Annotation of Human Serum Glycans That Yields Biological Variations. *Anal. Chem.* **87**, 7754–7762 (2015).
 300. Harvey, D. J. *et al.* Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. *Proteomics* **9**, 3796–3801 (2009).
 301. Storey, J. D. A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 479–498 (2002).
 302. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9440–9445 (2003).
 303. Cleveland, W. S. Robust locally weighted regression and smoothing

- scatterplots. *J. Am. Stat. Assoc.* **74**, 829–836 (1979).
304. Wang, J. *et al.* Fc-glycosylation of IgG1 is modulated by B-cell stimuli. *Mol. Cell. Proteomics* **10**, 1–12 (2011).
 305. Mackiewicz, A., Ganapathi, M. K., Schultz, D. & Kushner, I. Monokines regulate glycosylation of acute-phase proteins. *J. Exp. Med.* **166**, 253–258 (1987).
 306. Colombo, M. *et al.* Quantitative levels of serum N-glycans in type 1 diabetes and their association with kidney disease. *Glycobiology* **31**, 613–623 (2021).
 307. Wu, Y. *et al.* Probing serum N-glycan patterns for rapid and precise detection of Crohn's disease. *Chem. Commun.* **57**, 11362–11365 (2021).
 308. Ozcan, S. *et al.* Serum glycan signatures of gastric cancer. *Cancer Prev. Res.* **7**, 226–235 (2014).
 309. Kim, J. H. *et al.* Mass spectrometric screening of ovarian cancer with serum glycans. *Dis. Markers* **2014**, 1–9 (2014).
 310. Comunale, M. A. *et al.* Total serum glycan analysis is superior to lectin-FLISA for the early detection of hepatocellular carcinoma. *PROTEOMICS – Clin. Appl.* **7**, 690–700 (2013).
 311. Pfeiffer, R. M., Webb-Vargas, Y., Wheeler, W. & Gail, M. H. Proportion of U.S. Trends in breast cancer incidence attributable to long-term changes in risk factor distributions. *Cancer Epidemiol. Biomarkers Prev.* **27**, 1214–1222 (2018).
 312. Krištić, J. *et al.* Glycans Are a Novel Biomarker of Chronological and Biological Ages. *Journals Gerontol. Ser. A* **69**, 779–789 (2014).
 313. Knežević, A. *et al.* Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* **20**, 959–969 (2010).
 314. Ding, N. *et al.* Human serum N-glycan profiles are age and sex dependent. *Age Ageing* **40**, 568–575 (2011).
 315. Ogden, C. L. *et al.* Trends in obesity prevalence by race and hispanic origin - 1999-2000 to 2017-2018. *JAMA* **324**, 1208–1210 (2020).
 316. Ellulu, M. S., Patimah, I., Khaza'ai, H., Rahmat, A. & Abed, Y. Obesity and inflammation: the linking mechanism and the complications. *Arch. Med. Sci.* **13**, 851–863 (2017).
 317. Clerc, F. *et al.* Human plasma protein N-glycosylation. *Glycoconj. J.* **33**, 309–343 (2016).
 318. Chen, Z., Huang, J. & Li, L. Recent advances in mass spectrometry (MS)-based glycoproteomics in complex biological samples. *Trends Anal. Chem.* **118**, 880–892 (2019).
 319. Holst, S. *et al.* Linkage-Specific in Situ Sialic Acid Derivatization for N-Glycan Mass Spectrometry Imaging of Formalin-Fixed Paraffin-Embedded Tissues. *Anal. Chem.* **88**, 5904–5913 (2016).
 320. Clift, C. L., Mehta, A. S., Drake, R. R. & Angel, P. M. Multiplexed imaging mass spectrometry of histological staining, N-glycan and extracellular matrix from one tissue section: a tool for fibrosis research. in *Multiplexed*

Imaging: Methods and Protocols In Press (2020).

321. Lyons, M. J., Poole, S. F., Brownson, R. C. & Lyn, R. Place Is Power: Investing in Communities as a Systemic Leverage Point to Reduce Breast Cancer Disparities by Race. *Int. J. Environ. Res. Public Heal.* 2022, Vol. 19, Page 632 **19**, 632 (2022).