

The Host Glycome in Health and Immunity

by

Rui Qin

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
University of Alberta

© Rui Qin, 2022

Abstract

Glycosylation is the enzymatic modification of biomacromolecules (proteins, etc.) with carbohydrates known as glycans. Naturally occurring in all organisms from viruses to animals, glycans have multifaceted functions in all aspects of biology. Aberrant glycosylation is often observed in diseases such as cancers and in developmental abnormalities. In infectious disease, studies have heavily focused on the glycans of the pathogens, which can determine the immunogenicity and virulence of the pathogens. The less studied host glycans are also critical in immunity, influencing antibody function, complement system activation, and modulate immune cell functions. Due to the structural diversity of glycans, pinpointing the host glycan subtypes that undergo changes during infections can be a challenging task. In the original work described in this thesis, high-throughput lectin microarray technology was utilized to profile the glycomes related to human health and infection (i.e., the repertoires of glycans) in multiple systems. Glycomic analyses revealed differential glycosylation in the contexts of (i) sex, age, BMI, (ii) Coronavirus disease 2019 (COVID-19) and (iii) vaccination, which are associated with immune response or disease outcome. Moreover, glycoproteomic analyses identified potential target glycoproteins that may contribute to the differences in glycosylation and immunity. Overall, these discoveries provide fresh insights into the roles of host glycans in infectious diseases and have important implications in the development of glycosylation-focused therapeutics.

Preface

(Mandatory due to collaborative work)

Research projects described in Chapter 2, 3 and 4 of this thesis are collaborative projects. My contributions include: Chapter 2 – sample processing, glycomics, and data analysis; Chapter 3 – plasma/animal tissue sample processing, plasma/animal tissue glycomics, glycoproteomics, and data analysis; Chapter 4 – sample processing, glycomics, glycoproteomics, and data analysis.

Research described in Chapter 2 of this thesis has not been made publicly available. Research described in Chapter 3 of this thesis has been posted on the preprint server medRxiv (doi: 10.1101/2022.06.06.22275981). Research described in Chapter 4 of this thesis has been published in *Journal of Proteome Research* (doi: 10.1021/acs.jproteome.2c00251).

Dedication

Dedicated to my family, mentors, and friends

Acknowledgments

First, I would like to thank my supervisor Prof. Lara K. Mahal for her tremendous guidance and support in my projects, without which it would not be possible for me to come this far in my graduate program. Professor Mahal has also been giving me so much encouragement that inspired me to pursue what I wish to achieve in and outside my professional life.

I would also like to thank my current and former colleagues in the lab. Every member of the lab helped created a lab environment that made any advances in my research project possible. Spending time with my colleagues has also been an enjoyable experience that I will cherish. I also thank my friends from other labs and departments, from whom I have received valuable advice on science and life.

I would like to acknowledge all my collaborators for giving me so much help and inspiration and completing so many remarkable research projects together. I also acknowledge my committee members and examiners.

Finally, I would like to thank my parents, my family and friends for their unconditional support in the challenging times during my graduate study.

Table of Contents

Abstract	ii
Preface	iii
Dedication	iv
Acknowledgments	v
List of Tables	ix
List of Figures or Illustrations	x
List of Symbols	xiii
Chapter 1. Glycosylation in Mammalian Immunity	1
1.1 Overview of mammalian protein glycosylation	1
1.2 Host glycosylation in innate immunity	3
1.3 Host glycosylation in adaptive immunity	7
1.4 Changes of host glycosylation in infectious diseases	8
1.5 Thesis aims	9
Chapter 2. Defining the Serum Glycomic Profiles of a General Population	11
2.1 Introduction	11
2.1.1 The serum glycoproteome in immunity	11
2.1.2 Lectin microarray as a high throughput glycome-profiling tool	12
2.2 Results and Discussion	15
2.2.1 Global variability of serum glycomes	15
2.2.2 Serum protein glycosylation is associated with gender, age and BMI	20
2.2.3 Glycoproteomic analysis of lectin enriched serum	26

2.3 Conclusion	29
2.4 Methods	30
Chapter 3. Discovery of Glycan Biomarkers in COVID-19	35
3.1 Introduction	35
3.2 Results and Discussion	37
3.2.1 α 2,6-sialylation is increased in severe COVID-19 patient plasma	37
3.2.2 α 2,6-sialylation and its biosynthesis are increased in post-mortem COVID-19 patient tissues	41
3.2.3 Differential α 2,6-sialylation of complement proteins is associated with COVID-19 severity	47
3.2.4 α 2,6-sialylation is induced by SARS-CoV-2 infection in ferret lungs	58
3.3 Conclusions	61
3.4 Methods	62
Chapter 4. Discovery of Glycan Biomarkers Predicting Vaccine Responses	71
4.1 Introduction	71
4.2 Results and Discussion	74
4.2.1 Pre-vaccination serum Le ^a differentiates high and non-responders	74
4.2.2 Serum Le ^a is enriched in complement proteins	81
4.2.3 Potential biological links between Le ^a and vaccine response	86
4.2.4 Post-vaccination glycomic changes in high responders indicate active immune response	89
4.3 Conclusions	93

4.4 Methods	94
References	102
Appendices	129

List of Tables

Table 1.1. Functional Roles of Host Glycosylation in Innate Immune System

Table 2.1. Descriptive Information of the Cohort

Table 2.3. Serum glycoproteins enriched by AIA.

Table 2.4. Serum glycoproteins enriched by diCBM40.

Table 3.1. Descriptive characteristics of the plasma cohorts.

Table 3.2 Descriptive characteristics of the patients of the autopsy tissue cohorts.

Table 3.3 List of SNA-enriched proteins in the pooled severe and mild COVID-19 plasma

Table 4.1. Associations of host glycosylation and vaccine response in influenza-vaccinated human cohorts from past and present studies.

Table 4.2. Characteristics of Study Participants.

Table 4.3. Differences in pre-vaccination microarray data of Lea-binding lectins between high responders and non-responders to each influenza strain.

Table 4.4. Serum glycoproteins enriched by BambL and/or anti-Le^a

List of Figures or Illustrations

Figure 1.1. Common subclasses of *N*- and *O*-linked glycans of mammalian glycoproteins.

Figure 1.2. The complement cascade.

Figure 2.1. Scheme of a simplified workflow of dual-color lectin microarray experiment.

Figure 2.2. Heatmap of the plasma glycomic data of the cohort generated with lectin microarray and clusters of lectins with annotations of the rough specificities.

Figure 2.3. Lectins ranked by median absolute deviation shown in a polar coordinate.

Figure 2.4. Volcano plot comparing lectin microarray data of samples from male and female participants.

Figure 2.5. Box plots comparing select lectin microarray data of samples from normal, overweight, and obese participants.

Figure 2.6. Box plots comparing select lectin microarray data of samples from participants aged 20-40, 40-60, and > 60 years.

Figure 3.1. Plasma glycomic profiles of COVID-19 positive and negative cohorts.

Figure 3.2. Boxplot analysis of select α 2,3-sialic acid-binding lectins and core 1/3 *O*-glycan-binding lectins between patient groups.

Figure 3.3. Glycosylation patterns in tissues (heart, kidney, liver, upper and lower lobe lung) of deceased COVID-19 patients and negative controls.

Figure 3.4. Representative images of lectin fluorescence staining (red) and nuclei staining (blue) in tissues of deceased COVID-19 patients and negative controls.

Figure 3.5. Representative images of ST6GAL1 and hematoxylin staining in the upper lobe lungs, lower lobe lungs and livers of deceased COVID-19 patients and negative controls.

Figure 3.6. Glycoproteomic analysis of α 2,6-sialic acid-containing proteins from mild and severe COVID-19 plasma.

Figure 3.7. Differential α 2,6-sialylation of complement C5 and C9 in three patient groups.

Figure 3.8. Representative images of C5b, C9 and nuclei staining in lower lobe lungs, upper lobe lungs and livers of deceased COVID-19 patients and negative controls.

Figure 3.9. Representative images of IL-6, CD163, vimentin and nuclei staining in the tissues of deceased COVID-19 patients and negative controls.

Figure 3.10. Heatmap of lung tissue glycomic profiles of SARS-CoV-2-infected ferrets and negative cohorts.

Figure 3.11. Time course changes in α 2,6-sialylation and immune response marker expression in ferret lung tissues.

Figure 4.1. Heatmap of lectin microarray data for d0 serum samples. Columns represent the participants and rows represent the probes.

Figure 4.2. Volcano plot comparing lectin microarray data for high responders and non-responders pre-vaccination; Partial biosynthetic routes of Lewis A antigen and type I blood group H (O) antigen.

Figure 4.3. Volcano plot comparing pre-vaccination lectin microarray data for high responders and non-responders, re-classified using the MLR-adjusted scores with estimated impacts of multiple confounding factors removed.

Figure 4.4. Volcano plot comparing lectin microarray data for high responders and low/moderate-responders pre-vaccination.

Figure 4.5. Scheme of the experimental approach of glycoproteomic analysis; Number of glycoproteins identified in BambL/anti-Le^a pulldown experiments; Gene ontology pathway enrichment analysis for glycoproteins enriched with BambL/anti-Le^a; Differential C4BP glycosylation.

Figure 4.6. A possible mechanism of how Lea levels impacts vaccine response.

Figure 4.7. Heatmap of lectin microarray data for post-vaccination (day 28) serum samples; Boxplots of Log₂ Fold-Change for paired d28 and d0 samples in non- and high-responders.







Figure 4.8. Boxplots comparing correlation coefficients between pre- and post-vaccination lectin microarray data.

Figure 4.9. Volcano plot comparing the fold differences (d28 vs. d0) lectin microarray data for high responders and non-responders.

Figure 4.10. Boxplots comparing post-vaccination microarray data of Lewis A-binding lectins.

List of Symbols

Symbol Nomenclature for Glycans (SNFG) used in this thesis:

<i>Symbol</i>	<i>Monosaccharide unit</i>
	Mannose
	Galactose
	Fucose
	<i>N</i> -Acetylglucosamine (GlcNAc)
	<i>N</i> -Acetylgalactosamine (GalNAc)
	<i>N</i> -Acetylneuraminic acid (sialic acid)

Chapter 1. Glycosylation in Mammalian Immunity

1.1 Overview of mammalian protein glycosylation

Glycosylation is the enzymatic modification of molecules with glycans, which are also known as oligosaccharides, carbohydrates or sugars. It is a ubiquitous phenomenon found in all organisms. Naturally glycosylated forms of proteins, lipids, nucleic acids and small molecules have been discovered.¹⁻³ Glycosylation has found significant roles in virtually all aspects of mammalian biology.^{1,4}

Studies of glycosylation in mammalian systems have primarily focused on glycoproteins.⁵ *N*-linked glycan (*N*-glycan, linked to asparagine) and *O*-linked glycan (*O*-glycan, linked to serine or threonine) are the two major types of glycan in mammalian glycoproteins. **Figure 1.1** lays out some common subclasses of *N*- and *O*-glycans. Other *N*- and *O*-glycans such as paucimannose *N*-glycan, core 3/4 *O*-GalNAc glycans and *O*-xylose glycans are either low in abundance or restricted to certain tissues or proteins.^{6,7} Some structural motifs of glycan are more commonly known as blood group antigens (e.g., blood group antigen B, Lewis A antigen), which can occur in either *N*-linked or *O*-linked glycans.

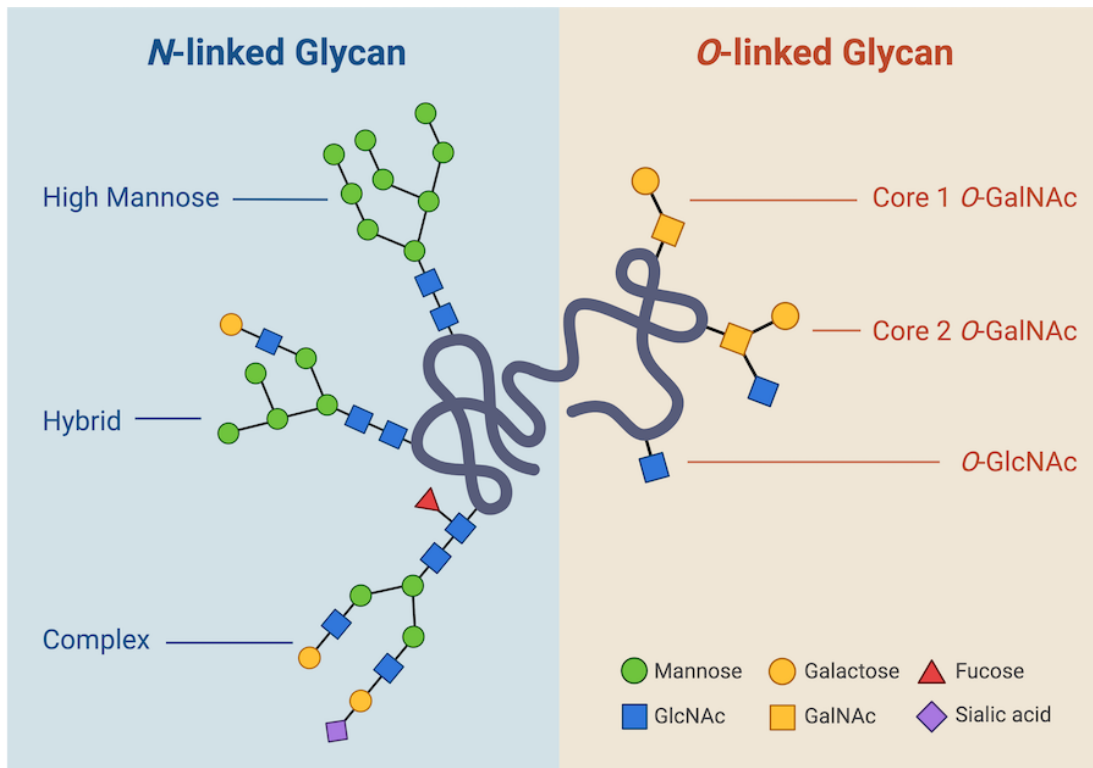


Figure 1.1. Common subclasses of *N*- and *O*-linked glycans of mammalian glycoproteins. Any depicted glycan is only an example of the glycans in the corresponding subclass.

The structures of glycans can significantly impact the physicochemical properties and biological activities of glycoconjugates.^{1,8} It is estimated that > 7,000 distinct glycan structures exist in mammals.^{9,10} The structural diversity of glycans arises from several factors including differences in 1) the composition of monosaccharides (the simplest carbohydrate unit for glycans biosynthesis), 2) the positions of the hydroxyl groups for glycosidic bond formation, 3) the α and β linkage of glycosidic bonds, and 4) branching.⁴ Despite the numerous possibilities of glycan structures, the glycan repertoire of a certain structural component (e.g., plasma, lung tissue) in an organism, simply known as the glycome, is maintained at equilibrium under normal circumstances via multi-level regulatory mechanisms that remain largely elusive.¹¹ Transcription

factors,¹²⁻¹⁴ epigenetic regulators,¹⁵ microRNAs,¹⁶ substrate availability,¹⁷ enzyme localization and transporting machineries¹¹ all have roles to play in the regulation of glycosylation. Changes in glycomes have been observed in a broad range of diseases including infectious diseases,¹⁸ cancers,^{4,19} autoimmune diseases^{20,21} and metabolic diseases.²² These changes are often mechanistically linked to pathogenesis.²³

This thesis is mainly focused on the intersection of glycosylation and immunity in human health. In the following sections of this introductory chapter, I first briefly explain the multifaceted roles of glycans in both innate and adaptive immunity, with a focus on host glycans. Next, I summarize the current state of the study of how host glycosylation changes in the context of infectious diseases. Finally, I point out the gaps in the current knowledge of host glycosylation in immunity and how my work would contribute to filling those gaps.

In this thesis, the Symbol Nomenclature for Glycans (SNFG, see List of Symbols) is used for depicting glycans.²⁴ To avoid excessive repetitions, glycosylation and related terms (glycan, glycome, etc.) are discussed in the context of mammalian glycoproteins hereinafter unless noted otherwise.

1.2 Host glycosylation in innate immunity

Innate immunity is the first-line defense mechanism against pathogens. The innate immune response features rapid action and broad specificity against pathogens. This is achieved by a set of receptors on stand-by mode that recognizes pathogen-associated molecular patterns (PAMPs). PAMP recognition initiates signaling pathways

that eventually lead to pathogen clearance, lysis of affected cells, inflammatory response, activation of adaptive immune response, etc.²⁵ Most PAMPs are the structural components of pathogens, including glycans.^{25,26} The study of immunogenic glycans of pathogens has a history almost a century long and has yielded remarkable results such as the invention of the Pneumococcal polysaccharide vaccine.²⁷ Comparatively, studying host glycosylation in innate immunity is a more recent event. In **Table 1.1**, I summarize the known roles host glycosylation plays in the innate immune system.

Table 1.1. Functional Roles of Host Glycosylation in Innate Immune System

Glycoform / Glycan Motif	Functional Role	Examples
<i>N</i> -glycan	Mutation of <i>N</i> -glycosylation sites reduces ligand affinity	Toll-like receptor 4 (TLR4), ²⁸ Natural Killer Cell Receptor 2B4 ²⁹
<i>N</i> -glycan	Removal of <i>N</i> -glycans impairs downstream signaling	TLR3 ³⁰ , C5a desArg ³¹
<i>N</i> -glycan	Removal of <i>N</i> -glycans changes complement system activity	CD46 ³² , C9 ³³
Complex <i>N</i> -glycan	Complex <i>N</i> -glycoforms are essential for forming active oligomers	Natural Killer Cell Receptor Nkp30 ³⁴
Branched <i>N</i> -glycan	Highly branched <i>N</i> -glycans promotes endocytosis of innate immune receptors	TLR2 and TLR4 ³⁵
Sialic acid	Lack of sialylation reduces ligand affinity	Macrophage mannose receptor (MMR) ³⁶
Sialic acid	Hypersialylation dampens pathogen-killing activity mediated by the receptor	TLR4 ^{37,38}
Sialic acid	Host sialic acid binding prevents complement activation	Factor H ³⁹

Sulfated glycosaminoglycan (GAG)	High level of GAG sulfation is required for innate immune lectin binding	Factor H variant ⁴⁰
O-glycan	O-glycan is required for ligand binding	Natural Killer Cell Receptor Ncr1 ⁴¹

The complement system is an essential part of the innate immune system (also the adaptive immune system). It is a humoral immune mechanism involving dozens of complement components, most of which exist in circulation.^{42,43} The activation of complement is a major immune event during the acute phase of infection. Complement is activated by a string of reactions known as the complement cascade, as illustrated in **Figure 1.2**. In the classical pathway of complement activation, antibodies bound to antigens form complexes with C1, triggering the reaction cascade. The lectin-mediated complement activation is triggered by glycan binding to a group of lectins (known as innate immune lectins or innate lectins), including the mannose-binding lectin (MBL or MBL2), ficolins and collectins. For example, MBL can recognize the high mannose glycans present on the surfaces of microorganisms such as *Neisseria meningitidis*.⁴⁴ The alternative pathway is initiated by the spontaneous hydrolysis of C3. Because most complement proteins are glycosylated, it can be postulated that their functions are impacted by glycans.⁴⁵ However, this has been experimentally validated only for some complement system proteins such as C5a (**Table 1.1**).

Immunity mediated by innate immune cells (mast cells, dendritic cells, natural killer cells, etc.) is another important arm of the innate immune system.²⁵ On the membranes of innate immune cells reside a collection of PAMP-recognizing receptors

such as Toll-like receptors (TLRs) and some C-type lectins (CTLs). Glycosylation of these receptors has also been found functionally significant, as exemplified in TLR4, MMR, etc. (Table 1.1).

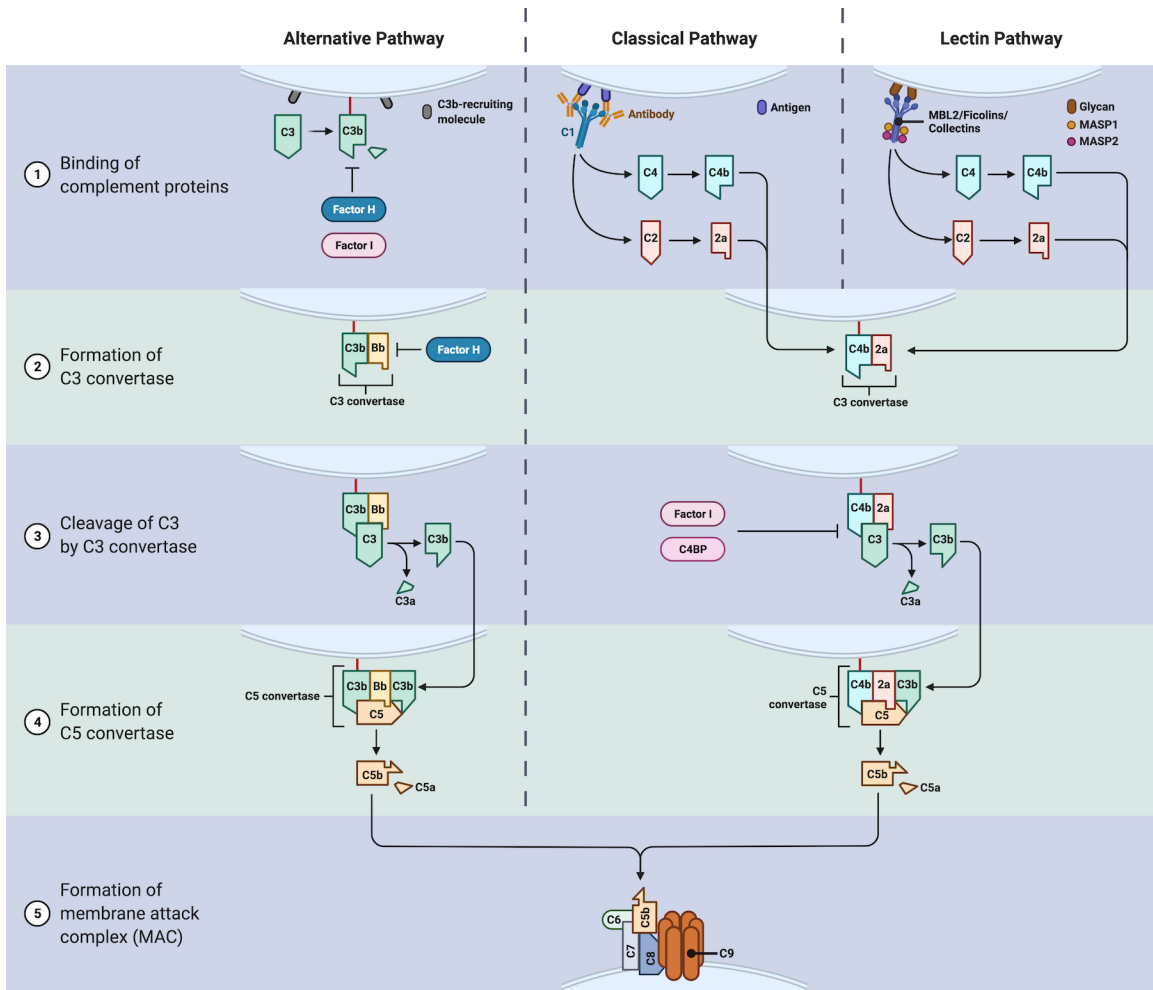


Figure 1.2. The complement cascade. Most of the proteins in the cascade are glycoproteins. Factor H, Factor I and C4BP are regulatory glycoproteins that inhibit the formation of the proteins/protein complexes indicated.

1.3 Host glycosylation in adaptive immunity

Adaptive immune system is characterized by high adaptability, high specificity, and the ability to develop immune memory.²⁵ Glycosylation of host proteins also have significant roles in adaptive immunity.

The antibody is perhaps one of the best-known proteins in adaptive immunity. Because of the importance of antibodies, the glycosylation of antibodies has been studied in great depth. All antibodies are glycoproteins, and the glycoforms of antibodies have significant impacts on their function and stability in circulation.^{8,46} Core fucosylation of IgG Fc N-glycan lowers its affinity to antibody Fc receptors, dampening the effector function of IgG. Sialylation of IgG increases its half-life in blood by blocking the interaction of IgG with scavenger receptors.^{47,48} In contrast, IgG bearing high mannose glycans are rapidly cleared from blood.⁴⁹ More recently, glycosylation of the IgG Fab chains was discovered to influence antigen binding and immune complex formation.⁵⁰ Functional roles of glycans have also been identified in IgA, IgM and IgE.^{21,51–54}

Glycosylation also has profound influence on cell-mediated adaptive immunity, which heavily relies on the glycosylated receptors on immune cell surfaces.⁵⁵ For example, core fucosylation of PD-1, an immunosuppressive T cell surface receptor, has a downregulatory effect on T cell activation. Binding of PD-L1, the endogenous ligand of PD-1, was diminished in T cells lacking surface core fucose.⁵⁶ Another example is CD28, an indispensable receptor for the activation of effector T cells. It was recently discovered that host sialic acid can bind CD28 in competition with CD80 (the cognate ligand of CD28), which leads to inhibition of effector T cell activation.⁵⁷

Galectins, selectins and Siglecs are lectins that have important regulatory roles in both the innate and adaptive immune system.^{58–60} They are traditionally considered to function mainly through interactions with host glycans instead of PAMPs. For example, Siglec-2 (CD22) on B cell surface has an inhibitory effect on B cell signaling upon binding to host sialoproteins.⁶¹ Binding of galectin-3 to highly branched *N*-glycans forms a local lattice structure on T cell surfaces, preventing the T cell receptor from interacting with their ligands and thus inhibiting TCR signaling.^{62,63} Multiple reports have shown changes in host glycosylation directly impact immunity via altered binding of galectins, selectins and Siglecs, mostly observed in cancers.^{58,64–66}

1.4 Changes of host glycosylation in infectious diseases

As elaborated above, it is clear that changes in the host glycome alter the immune system function and thus can impact disease outcomes. Indeed, in infectious diseases, the glycome of the host undergoes changes associated with severity and altered immune response.¹⁸ Changes in glycosylation of IgG during infections have been the most studied. In COVID-19 (Coronavirus disease 2019), HIV (human immunodeficiency virus) infection and tuberculosis, fucosylation, galactosylation and/or lack of bisected GlcNAc of pathogen-specific IgG are consistently associated with poor outcomes.^{67–72} It is believed that those glycoforms render IgG ineffective in mediating antibody-dependent cell-mediated cytotoxicity (ADCC). In another study among patients of bacterial sepsis, significant plasma IgG glycan hydrolysis was found in the most severe cases (septic shock).⁷³ Aside from antibody glycosylation, the Mahal Lab has

been exploring infection-associated changes in host glycosylation on a broader scope. Recently, they identified an association between the severity of influenza and high mannose glycans in lungs.⁷⁴ Higher levels of high mannose glycans may enhance mannose binding lectin-mediated complement activation, which promotes inflammatory responses.

Another interesting aspect of host glycosylation changes is found in enteric infection studies. By competing with invading pathogens, the gut microbiota provides their mammalian hosts with a special type of immunity.⁷⁵ Remodeling of host glycosylation has been found to be heavily involved in this immune mechanism. In mice, fucosylation of the gut epithelium is upregulated in response to infection by *Citrobacter rodentium*.⁷⁶ This benefits the growth of commensal bacteria that use fucosylated glycans as a source of nutrition. As the commensal bacteria community thrives, the pathogenic *Citrobacter rodentium* is suppressed. Similar mechanisms have been found in infections by *Enterococcus faecalis* and *Clostridium difficile*.^{77,78} These findings have important implications in inflammatory bowel disease (IBD), in which the gut epithelial glycosylation is often impaired.⁷⁹

1.5 Thesis aims

As previously discussed, our knowledge of how host glycosylation is remodeled in infectious diseases is currently scarce. This thesis presents my efforts to contribute to this area of study from three perspectives. First, I define the serum glycomes of a general population to gain insights into the links between glycosylation and differential

immune responses across populations (Chapter 2). Next, I investigate the host glycosylation changes in COVID-19 and how these changes are associated with severity, which points to a new route of treating severe infections (Chapter 3). In the final chapter (Chapter 4), I explore the association of host glycosylation with responses to influenza vaccination, the most important method to provide a population with immunity. Chapter 4 aims to be a starting point to fill the void of our understanding about the role of glycosylation in determining vaccination outcomes.

Chapter 2. Defining the Serum Glycomic Profiles of a General Population

2.1 Introduction

2.1.1 The serum glycoproteome in immunity

Blood, consisting of blood cells and serum/plasma, is one of the most important samples for monitoring the immune system in clinical practice. Glycoproteins make up around half of total serum proteins.⁸⁰ Many serum glycoproteins are closely associated with immunity, including antibodies (IgG, IgA, IgM, IgE and IgD), complement proteins, cytokines and acute phase proteins. Glycosylation is important to the functions of these serum proteins, and changes in glycosylation in these proteins have been implicated in disease (see Chapter 1). Besides influencing the functions of glycoproteins, changes in glycosylation also influence interactions of glycans with circulating immune-modulating lectins such as Siglecs and galectins. For example, sialylated IgM, a ligand of Siglec receptors of peripheral immune cells, is indispensable for suppressing immune cell activation in healthy individuals.^{53,81} Sialylated proteins also bind complement factor H, which downregulates the alternative pathway of complement activation.³⁹ It is therefore important to advance our understanding of immunity through the lens of serum protein glycosylation in both healthy and pathological states.

Benchmarking the baselines of biomolecules in populations with diverse demographic profiles is essential in immunology research and its clinical translation. Variations in biomolecules among the general populations have been associated with differential levels of immune responses. For example, different levels in sex hormones

in men and women influence the expression of immune receptors and cytokines, explaining some differences in immune responses in infection.⁸² Age also has profound influences on immunity. Aged people tend to have higher levels of pro-inflammatory, immunosuppressive cytokines in their blood and peripheral blood cells, mapping onto generally weaker immune responses.^{83,84}

Variations in glycosylation among general populations have been found in previous studies, summarized in a recent review.⁸⁵ While past work has shown the varying nature of the human serum/plasma glycome, it also has several limitations. First, they only focused on *N*-glycomes due to limitations of mass spectrometry-based techniques (see section 2.1.2 below). Secondly, as subsequent glycoproteome analyses were not performed, the identities of glycoproteins bearing the varying glycoforms were unknown. Third, linkages of glycans, which are important for the biological activities of glycans, were not resolved in some of those studies. In this chapter, I detail my work on lectin microarray-based, comprehensive serum glycomic and glycoproteomic analyses in a large (n = 352) human cohort. This work provided not only new data on human glycome variation associated with gender, age and BMI, but also fresh insights into how glycosylation may impact immunity.

2.1.2 Lectin microarray as a high throughput glycome-profiling tool

Given the structural diversity of glycans, pinpointing specific glycans or glycan motifs that change between groups of complex biological samples (e.g., serum/plasma, tissue, mucus) can be an arduous task. Therefore, high throughput glycomic

technologies that profile multiple glycan structures simultaneously are invaluable for the study of glycosylation. Lectin microarrays are one of the most widely used glycomic profiling technologies. Lectins are carbohydrate-binding proteins naturally existing in almost all organisms from viruses and bacteria to plants and mammals. Lectins bind structurally different glycan motifs.⁸⁶ Lectin microarrays utilize a collection of lectins, immobilized on solid surfaces, to probe differences in certain glycan motifs among sample groups. Current lectin microarrays have expanded to include engineered lectins and anti-glycan antibodies. A scheme of lectin microarray workflow is shown in **Figure 2.1**. In brief, glycoproteins are tagged with fluorescent dyes and hybridized with the microarrays. Fluorescence intensities (for single-color experiment) or fluorescence intensity ratios (for dual-color experiment) are used to quantitate the glycan motifs corresponding to the specificities of the lectins.

Compared to other two predominant glycomic technologies, liquid chromatography (LC) and mass spectrometry (MS), lectin microarrays have several advantages.⁷ First, the workflow of lectin microarrays is simpler. Glycan release is generally not needed in lectin microarray-based glycomic profiling. For LC and MS, glycans must be released from glycoproteins and purified prior to analysis. Related to this, LC- and MS-based glycomic analyses have been heavily biased towards the *N*-glycomes. The enzyme PNGase F (Peptide:*N*-glycosidase F) can be used to cleave all *N*-glycans from mammalian glycoproteins with high efficiency. However, the counterpart of PNGase F for *O*-glycan has not been discovered. In addition, LC/MS databases for *O*-glycans were lacking. These factors make LC- and MS-based *O*-glycomic analysis a

challenging task. Secondly, by using linkage-specific lectins, lectin microarrays readily distinguish glycan isomers with different linkages. For LC and MS, however, differentiating glycan isomers is more challenging since they have the same molecular mass and are usually difficult to separate by chromatography. Third, probes used in lectin microarray can be directly used in downstream histochemical and glycoproteomic studies, as exemplified later in this chapter and in the subsequent chapters. Lectin microarrays have been used in a wide variety of studies including analysis of exosomes, host-response to influenza, and identification of cancer drivers in human tissues.^{74,87–92}

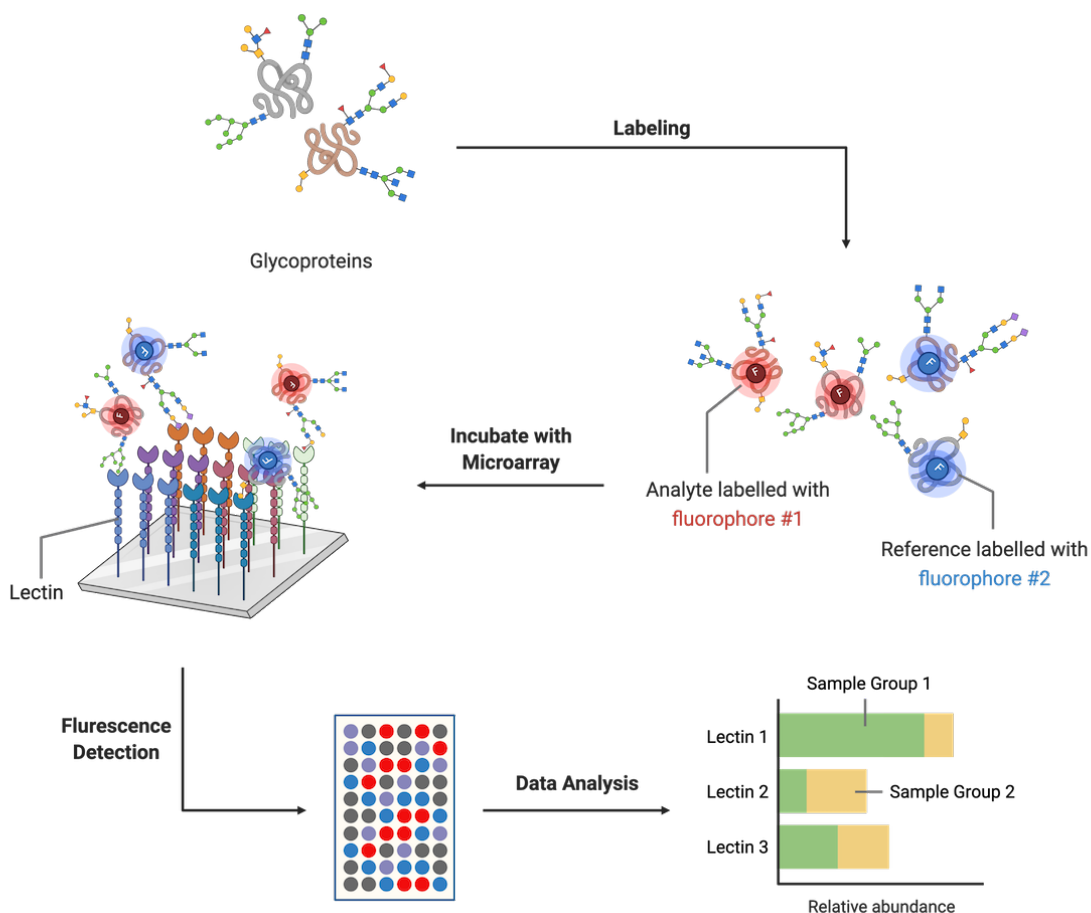


Figure 2.1. Scheme of a simplified workflow of dual-color lectin microarray experiment. Fabrication process of lectin microarray slides is not shown.

2.2 Results and Discussion

2.2.1 Global variability of serum glycomes

Serum glycomes from a non-disease-focused cohort (general population) of 352 individuals (**Table 2.1**) were analyzed using lectin microarray. The probes (see Appendix I) contained well characterized lectins covering a wide range of glycan motifs, as well as several immunoglobulin-binding proteins (protein A, protein G and protein L) and antibodies against some endogenous, immune-modulating serum lectins. Serum contains a diverse array of glycans, as seen by the fact that 81 of 84 lectins gave significant signals. A heatmap overview of the glycomic data of the cohort is presented in **Figure 2.2**.

Table 2.1. Descriptive Information of the Cohort

Gender (n)	Male: 220; Female: 130
Age Ranges (n)	20-30 yrs: 62; 30-40 yrs: 104; 40-50 yrs: 83; 50-60 yrs: 58; 60-70 yrs: 38; > 70 yrs: 15
Median BMI (IQR)	24.4 (22.3 - 27.2)
BMI Category (n)	Normal ($18 \leq \text{BMI} < 25$): 164; Overweight ($25 \leq \text{BMI} < 30$): 97; Obese ($30 \leq \text{BMI}$): 37;
Ethnicity (n)	Asian: 96; Caucasian: 237; Other: 19

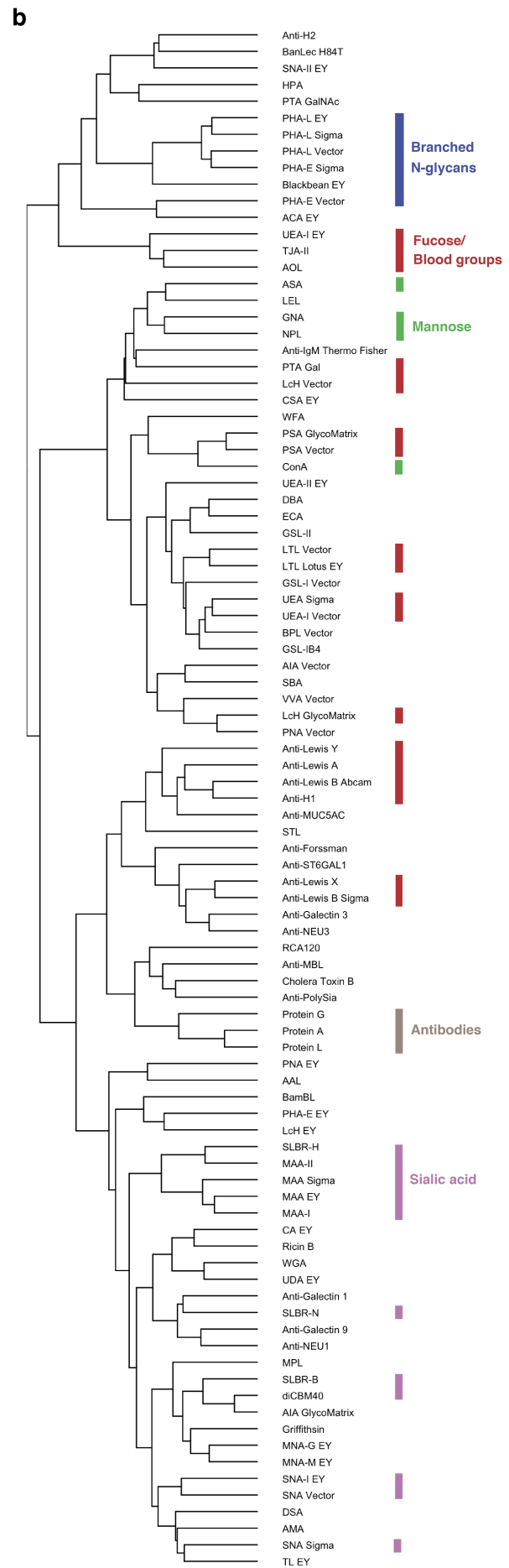
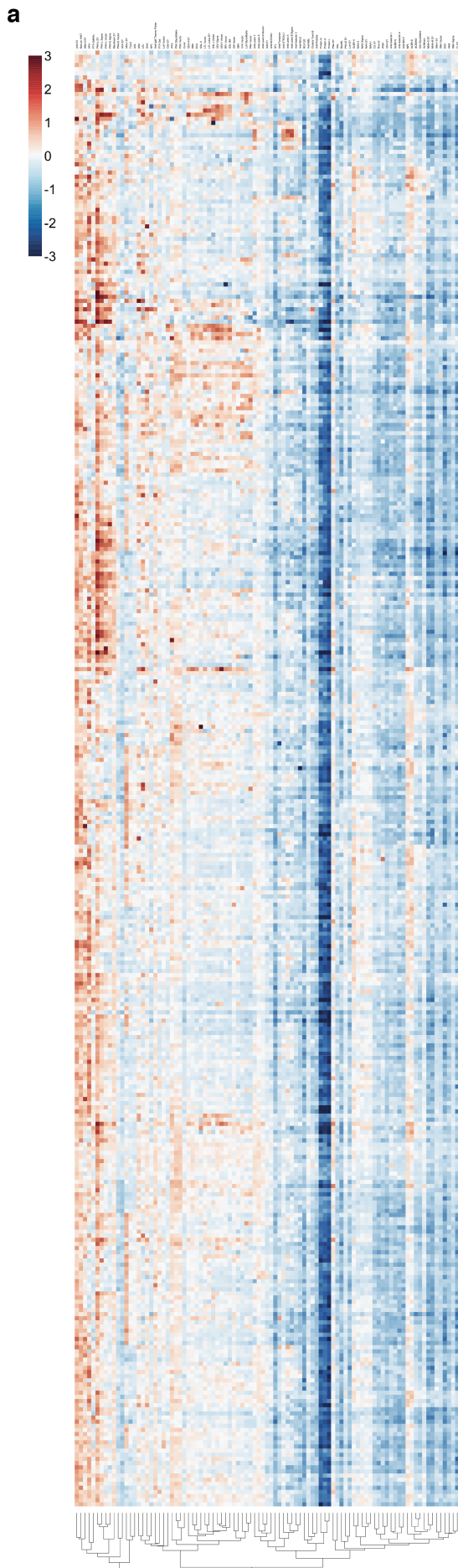


Figure 2.2. (a) Heatmap of the plasma glycomic data of the cohort generated with lectin microarray. Data were hierarchically clustered using Pearson distance. **(b)** Clusters of lectins in (a) with annotations of the rough specificities.

To identify the most varying glycan motif in human sera, median absolute deviation of the microarray data for each lectin was computed. Groups of probes of similar specificities were found within the top quartile deviations, including probes specific for β 1,6-branched glycans (PHA-L), *O*-GalNAc glycans (HPA, PNA, SLBR-B, SLBR-H), type II blood group antigen H (UEA-I, PTA GalNAc, anti-H2), and immunoglobulins (protein A, protein G, protein L) (**Figure 2.3**). My analysis of glycan motif variation mapped onto previous knowledge. PHA-L binds branched *N*-glycans with β 1,6-GlcNAc linkage, synthesized by the enzyme MGAT5 (alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A). β 1,6-branched glycan plays important roles in regulation of T cell-mediated immunity.⁹³ It prevents prolonged activation of effector T cells and maintains the function of regulatory T cells.^{94,95} *Mgat5* knockout mice developed autoimmune diseases.^{93,96} A GWAS (genome-wide association study) study analyzing 3533 individuals found one of the loci explaining the variations of the plasma *N*-glycomes was located in *MGAT5*, indicating a significant impact of MGAT5 on plasma glycan variation compared to many other glycoenes.⁹⁷ It was later found that *MGAT5* polymorphism was indeed associated with *MGAT5* expression.⁹⁸ The lectin microarray data provided additional evidence that β 1,6-branched glycans are one of the most variable glycan signatures in human serum. This glycan may be a promising indicator of the susceptibility to autoimmune diseases.

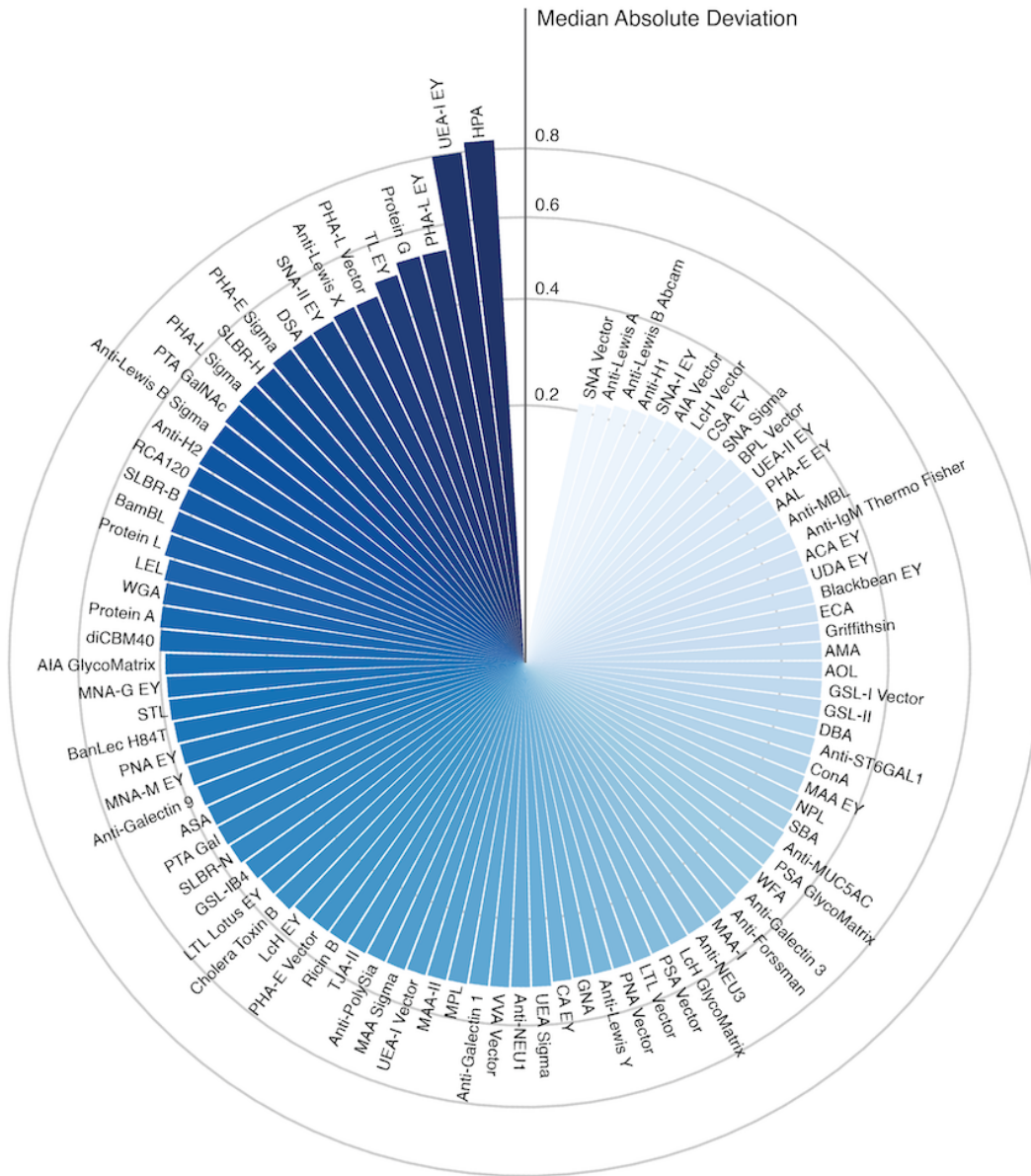


Figure 2.3. Lectins ranked by median absolute deviation shown in a polar coordinate.

Some of the most highly variable probes bind to blood group antigens, an expected result given the diversity of blood group antigens in the human population.⁹⁹ HPA binds motifs with terminal α GalNAc such as blood group antigen A and Tn antigen, often found in *O*-glycans.⁸⁶ In the context of plasma, they most likely bind to blood group

antigen A, as Tn antigen is rarely detected in healthy tissues including plasma.^{100,101} In the present study, the most varying blood group antigens identified seemed to be blood group antigen H (the antigen that determines the commonly known blood group O) and blood group antigen A, consistent with the fact that blood type O and blood type A are the most prevalent ABO blood types, consisting roughly 28-67% and 19-40% of the population, respectively.⁹⁹ Blood group antigens impact host susceptibility to pathogens and disease severity.^{99,102}

An interesting finding is the high variation of sera glycoprotein binding to PNA, SLBR-H and SLBR-B. PNA recognizes Gal β 1,3-GalNAc structure. Additionally, SLBR-H/SLBR-B can bind α 2,3-sialylated Gal β 1,3-GalNAc.^{86,103} These together hinted at a high variation of the Gal β 1,3-GalNAc core, which is synthesized by the enzyme C1GALT1 (core 1 synthase). Gal β 1,3-GalNAc and C1GALT1 are associated with tumor progression and survival in cancer patients.^{104,105} The biological significance of differential expression of Gal β 1,3-GalNAc in general populations is unclear.

In contrast, the least varying glycan motif seemed to be α 2,6-sialic acid (SNA, SNA-I) and some type I LacNAc-containing blood group antigens (anti-H1, anti-Lewis A) (**Figure 2.3**). The relatively low variation of α 2,6-sialic acid may be due to tight regulation by the asialoglycoprotein receptor on hepatocytes.^{106,107} From the data, α 2,3-sialic acid (MAA-II, SLBR-B, SLBR-N, diCBM40) and polysialic acid / α 2,8-sialic acid (anti-PolySia) seemed more variable than α 2,6-sialic acid.

2.2.2 Serum protein glycosylation is associated with gender, age and BMI

To understand the variation of glycosylation with basic demographic profiles (gender, age, BMI) in this cohort, participants were grouped by these factors. Comparison of lectin microarray data between groups revealed widespread differences in serum glycosylation associated with these factors. Men had higher levels of (i) *O*-glycans (MPL, AIA, SLBR-B, SLBR-H), (ii) α 2,3-sialic acids (SLBR-B, SLBR-H, MAA-II, diCBM40), and type II blood group H antigen (UEA, UEA-I) (**Figure 2.4**) in their sera glycomes. When comparing between normal, overweight, and obese participants, some glycans seemed to trend with BMI, including (i) *O*-glycans (MPL, AIA, PNA, MNA-G), (ii) α 2,3-sialic acids (SLBR-B, SLBR-H, diCBM40), and (iii) α 2,6-sialic acids (SNA, diCBM40) (**Figure 2.5**). For the comparison of age, participants were divided into three groups: 20-40 years, 40-60 years, and > 60 years. Lectin binding to *O*-glycans (MPL, AIA, PNA), (ii) α 2,3-sialic acids (MAA, MAA-II, SLBR-B, diCBM40), and (iii) core fucose (PSA) showed some age-dependent changes (**Figure 2.6**).

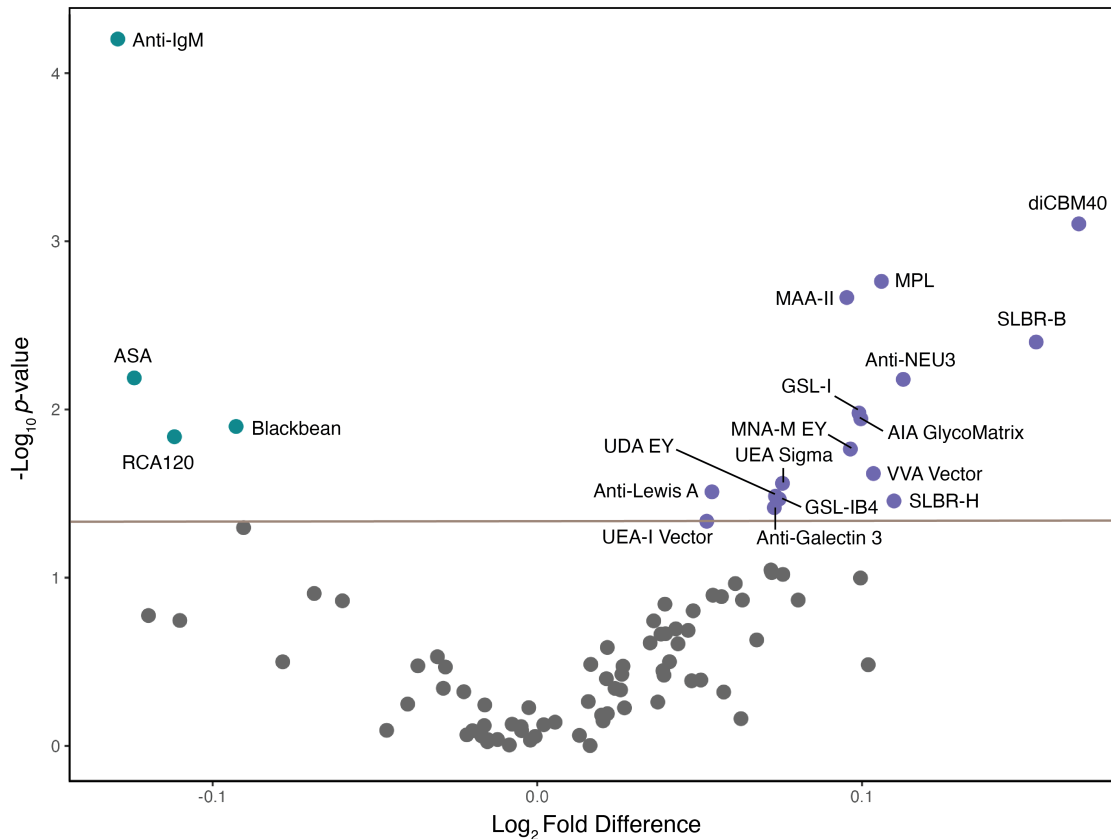


Figure 2.4. Volcano plot comparing lectin microarray data of samples from male ($n = 220$) and female ($n = 130$) participants. Mann–Whitney U test was used to determine p -values. Probes with $p < 0.05$ are labelled and colored in either violet (higher binding to male samples) or turquoise (lower binding to male samples).

In the univariate analysis above, *O*-glycans and α 2,3-sialic acids associated with all three factors examined. To unravel the effects of multiple factors, data from corresponding lectins (AIA, MPL, SLBR-B, SLBR-H, and diCBM40) were linearly modelled using sex, BMI and age as variables (**Table 2.2**). The results showed BMI had the greatest impact on to serum glycoprotein binding to all five modelled lectins, indicating strong association of metabolic status with serum glycomes. In contrast, age only had a minor reducing effect on MPL data, which was not replicated in AIA, which

has a similar specificity. Being male had a comparable effect as being obese/overweight on data from diCBM40, a pan-sialic acid binder with preference for α 2,3- over α 2,6-sialic acid.¹⁰⁸

The impact of obesity on sialic acid was not restricted to a certain type of sialic acid since the sera from the obese had greater binding to most of the sialic acid-binding lectins on the arrays. In concordance with this, an early study found total plasma sialic acid level positively correlated with higher body fat content in a diabetic cohort.¹⁰⁹ Obesity had a particularly high influence on SLBR-B and SLBR-H (**Figure 2.5; Table 2.2**), which both recognize α 2,3-sialic acids on *O*-glycans. The changes observed in SLBR-B and SLBR-H can be attributed to increased activity of α 2,3-sialylation machinery or increased sialylated *O*-glycoproteins. The former would likely implicate ST3GAL1 (CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,3-sialyltransferase 1) and ST3GAL2 (CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,3-sialyltransferase 2), the major enzymes that α 2,3-sialylate *O*-glycans. Evidence of altered activities of these enzymes in obesity is currently lacking.

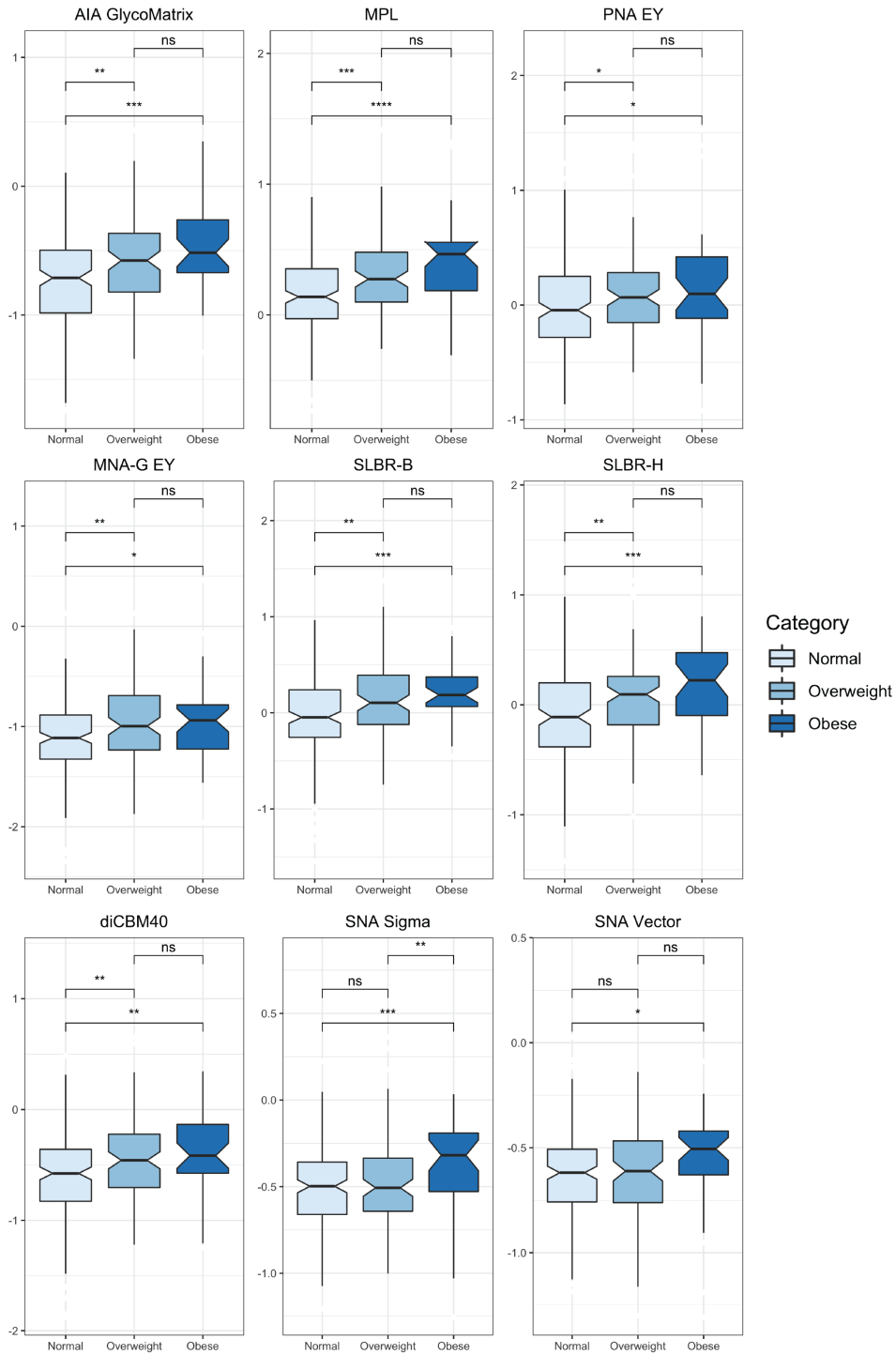


Figure 2.5. Box plots comparing select lectin microarray data of samples from normal (n = 164), overweight (n = 97), and obese (n = 37) participants. Mann–Whitney U test was used to determine p-values. *: p < 0.05; **: p < 0.01; ***: p < 0.001; ****: p < 0.0001; ns: p >= 0.05.

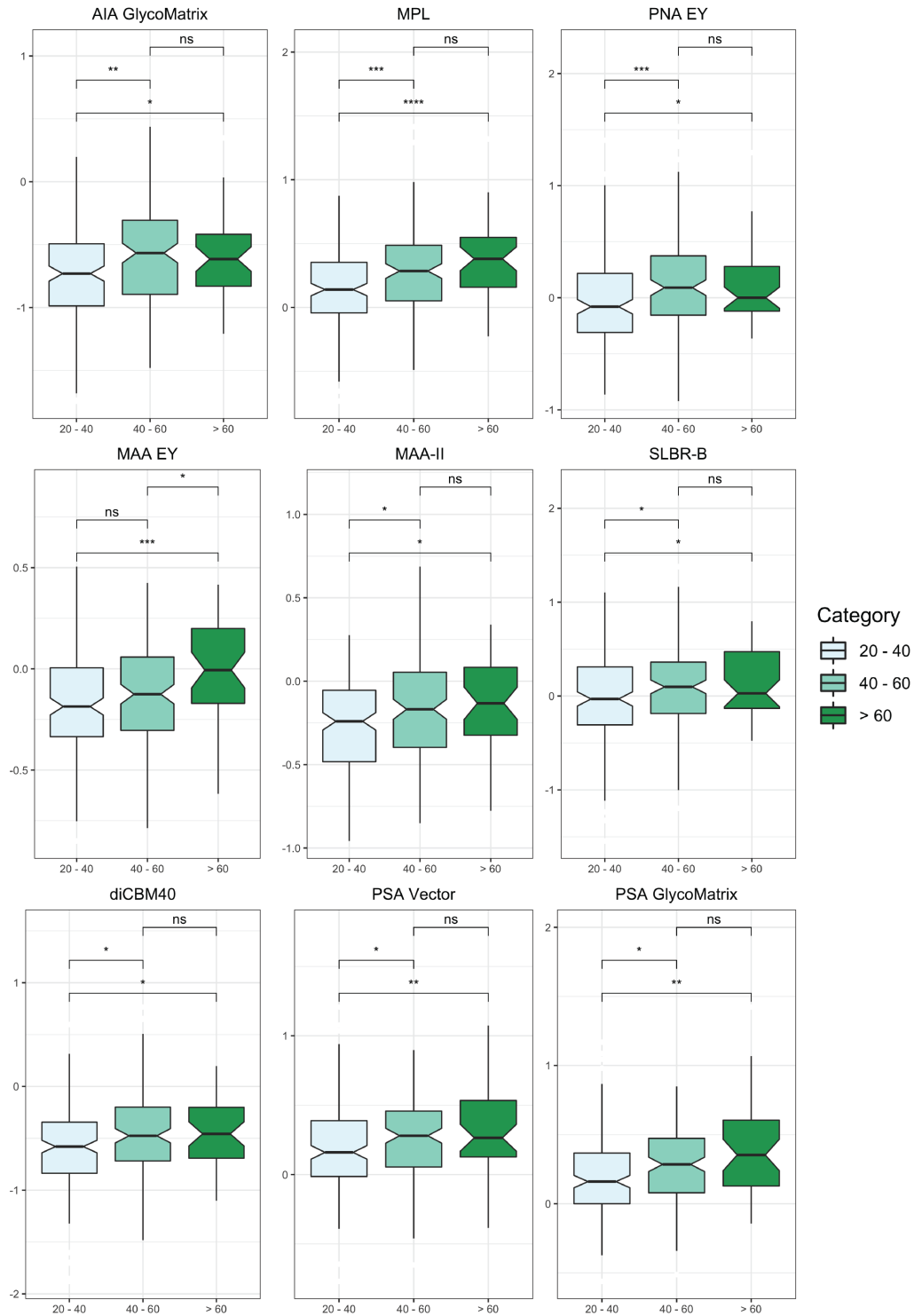


Figure 2.6. Box plots comparing select lectin microarray data of samples from participants aged 20-40 (n = 166), 40-60 (n = 141), and > 60 (n = 53) years. Mann–Whitney U test was used to determine p-values.

*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$; ns: $p \geq 0.05$.

Table 2.2. Key statistics of the linear models.

Lectin	Variable	Coefficient	Standard error	p-value
AIA GlycoMatrix	Male	0.074	0.045	0.099
	Obese	0.252	0.069	0.000
	Overweight	0.130	0.048	0.007
	Age Range 20 - 40	-0.044	0.070	0.535
	Age Range 40 - 60	-0.011	0.070	0.880
MPL	Male	0.060	0.037	0.104
	Obese	0.234	0.057	0.000
	Overweight	0.128	0.040	0.001
	Age Range 20 - 40	-0.141	0.058	0.015
	Age Range 40 - 60	-0.089	0.058	0.123
SLBR-B	Male	0.091	0.051	0.076
	Obese	0.282	0.079	0.000
	Overweight	0.191	0.055	0.001
	Age Range 20 - 40	-0.039	0.081	0.630
	Age Range 40 - 60	-0.010	0.080	0.898
SLBR-H	Male	0.060	0.051	0.240
	Obese	0.250	0.079	0.002
	Overweight	0.161	0.055	0.004
	Age Range 20 - 40	-0.043	0.080	0.595
	Age Range 40 - 60	-0.028	0.080	0.726
diCBM40	Male	0.129	0.047	0.007
	Obese	0.199	0.073	0.007
	Overweight	0.145	0.051	0.005
	Age Range 20 - 40	-0.037	0.074	0.620
	Age Range 40 - 60	-0.004	0.074	0.953

Note: *p*-values less than 0.05 are bolded.

2.2.3 Glycoproteomic analysis of lectin enriched serum

The analysis above showed the *O*-glycans and sialic acids (α 2,3-linked and α 2,6-linked) were the glycoforms most impacted by age, sex, and BMI. To identify the potential origin of glycosylation shifts in serum samples, glycoproteomic analysis was performed on pooled serum samples to identify the glycoprotein compositions. In brief, glycoproteins in the pooled sample were pulled down by bead-conjugated AIA and diCBM40 separately, and enriched samples were analyzed by mass spectrometry. AIA and diCBM40 were chosen because they have relatively broader specificities for *O*-glycans and sialic acids, and thus will likely enrich more protein targets.

AIA enriched 35 serum glycoproteins (**Table 2.3**) and diCBM40 enriched 20 (**Table 2.4**). As expected, the most enriched proteins by AIA included IgA, hemopexin, and α -2-HS-glycoprotein, all of which are among the most abundantly *O*-glycosylated proteins.¹⁰¹ Haptoglobin and haptoglobin-related protein, both binding hemoglobin,¹¹⁰ were among the top 3 enriched glycoproteins in diCBM40 pulldown. Haptoglobin, which can be induced in the white adipose tissue via obesity-associated cytokines,¹¹¹ are elevated in the sera of obese individuals compared to the non-obese,^{112,113} providing one explanation for the increased diCBM40 binding to the obese sera. However, serum levels of glycoproteins such as IgA and α -2-HS-glycoprotein has no clear association with obesity to date. IgA *O*-glycosylation changes in pathological conditions.²¹ Future studies could focus on whether the glycosylation status of IgA or other AIA/diCBM40-enriched serum glycoproteins also changes in obesity.

As for SLBR-B and SLBR-H, enrichment of plasma proteins with these two lectins has been reported in a previous study. They pulled down a small subset of *O*-glycosylated proteins in plasma, including PRG4 (proteoglycan 4), ITIH4 (inter-alpha-trypsin inhibitor heavy chain H4), and C1-INH (plasma protease C1 inhibitor).¹⁰³ PRG4 is an abundant, heavily *O*-glycosylated protein with roles in modulating immune responses and is known to increase in obesity.^{114–116} Therefore, higher levels of α 2,3-sialylated *O*-glycans observed in this study could simply be due to elevated PRG4 protein expression. However, since *O*-glycosylation plays an important role in PRG4 function and interaction with lectins,¹¹⁷ it could be worthwhile to study whether the glycosylation status of PRG4 is also changing in obesity, and whether it is associated with altered immunity.

Table 2.3. Serum glycoproteins enriched by AIA.

Uniprot Accession	Protein name	Average number of spectral matches
P01834	Immunoglobulin kappa constant	87.00
P01876	Immunoglobulin heavy constant alpha 1	68.67
P0DOY2	Immunoglobulin lambda constant 2	27.33
P02768	Serum albumin	19.00
B9A064	Immunoglobulin lambda-like polypeptide 5	18.00
P02790	Hemopexin	13.67
P0C0L5	Complement C4-B	11.00
P02765	Alpha-2-HS-glycoprotein	10.67
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	10.33
P01871	Immunoglobulin heavy constant mu	10.33
P04114	Apolipoprotein B-100	9.67
P01857	Immunoglobulin heavy constant gamma 1	9.00

P01023	Alpha-2-macroglobulin	8.67
P01024	Complement C3	8.33
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	7.67
P02647	Apolipoprotein A-I	5.67
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	5.00
P01860	Immunoglobulin heavy constant gamma 3	5.00
Q9NZT1	Calmodulin-like protein 5	4.67
P01861	Immunoglobulin heavy constant gamma 4	4.33
P01042	Kininogen-1	4.33
P00738	Haptoglobin	4.00
Q86YZ3	Hornerin	3.67
P00747	Plasminogen	3.67
P02751	Fibronectin	3.33
P01591	Immunoglobulin J chain	3.33
P10909	Clusterin	3.00
P00736	Complement C1r subcomponent	2.33
P02750	Leucine-rich alpha-2-glycoprotein	2.33
P47929	Galectin-7	2.33
P09871	Complement C1s subcomponent	2.00
Q08188	Protein-glutamine gamma-glutamyltransferase E	2.00
P01859	Immunoglobulin heavy constant gamma 2	2.00
P01009	Alpha-1-antitrypsin	1.67
P02760	Protein AMBP	1.33

Table 2.4. Serum glycoproteins enriched by diCBM40.

Uniprot Accession	Protein name	Average number of spectral matches
P00738	Haptoglobin	39.00
P01023	Alpha-2-macroglobulin	26.67
P00739	Haptoglobin-related protein	19.33
P01871	Immunoglobulin heavy constant mu	11.00
P01834	Immunoglobulin kappa constant	20.67
P01876	Immunoglobulin heavy constant alpha 1	12.33

P00751	Complement factor B	6.33
P00450	Ceruloplasmin	9.33
P10909	Clusterin	6.67
P01857	Immunoglobulin heavy constant gamma 1	7.33
P0DOY2	Immunoglobulin lambda constant 2	5.67
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	5.00
P01011	Alpha-1-antichymotrypsin	6.33
P02743	Serum amyloid P-component	2.67
P02790	Hemopexin	5.00
P0C0L4	Complement C4-A	3.33
P02768	Serum albumin	2.67
B9A064	Immunoglobulin lambda-like polypeptide 5	4.00
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	4.67
P02763	Alpha-1-acid glycoprotein 1	4.33

2.3 Conclusion

Serum/plasma glycosylation is influenced by multiple factors that can include genetics, sex/gender, age, BMI, disease state, physical exercise, and many others.^{18,85,118–123} Despite ample *N*-glycomic data in the past two decades, little efforts have been made to pursue the biological associations between glycan markers and immunity in general populations. In the present study, lectin microarray, which covers most of the *N*- and *O*-glycomes, was first used to probe the variations of serum glycome in a general population. Analysis of the microarray data identified multiple changes in serum glycosylation associated with gender, age, and BMI. Many of the changes were linked to *O*-glycans, which were often neglected in previous studies. Thus, this work has provided fresh insights into human glycome variation. The most notable finding from this

study is the strong association between obesity and increased *O*-glycans and sialic acids, especially α 2,3-linked sialic acids. Obesity can negatively impact immune responses to pathogens and vaccines through multiple mechanisms (cytokine secretion, immune cell recognition, etc.),^{124–126} in which *O*-glycans and sialic acids may play contributing roles (discussed in 2.2.3). Furthermore, glycoproteomic analysis identified glycoproteins candidates bearing these glycan motifs (e.g., IgA, haptoglobin). The glycosylation states of these proteins may change in obesity, which can be investigated in future studies.

2.4 Methods

Lectin Microarray

Total protein concentrations of serum samples were measured with *DC*TM protein assay (Bio-Rad Laboratories). To label glycoproteins, 10 μ g total protein of each sample was diluted in PBS to 27 μ L. Then 3 μ L of 1M NaHCO₃ (in water) and 0.21 μ L Alexa FluorTM 555 NHS ester (10 mg/mL in DMSO, Thermo Fisher Scientific) was added to and thoroughly mixed with the diluted sample. The vessel was wrapped in foil and incubated on a shaker at room temperature. After 1 hour, unconjugated dye molecules were removed with ZebaTM dye and biotin removal filter plates (Thermo Fisher Scientific). The reference standard, NIST human serum 909c (Millipore Sigma), was labelled with Alexa FluorTM 647 NHS ester (Thermo Fisher Scientific) using a similar protocol (i.e., amounts of reagents scaled linearly to total protein amount). After labelling, 10 μ g of each sample labelled with Alexa FluorTM 555 was mixed with 10 μ g of

reference standard labelled with Alexa Fluor™ 647. The volume of this dual-color mixture was adjusted to 50 µL with PBS, and then to 100 µL with 0.1% PBST.

Lectin microarray slides were fabricated as previously described in a published protocol.¹²⁷ The list of printed lectins can be found in the appendices. Each dual-color mixture was hybridized on the microarrays on a shaker at room temperature in the dark. After 1 hour, the microarrays were washed twice with 0.005% PBST for 10 minutes, once with PBS for 5 minutes, briefly rinsed once with deionized water, and finally dried by centrifugation. Fluorescence signals were obtained with Genepix™ 4400A fluorescence slide scanner (Molecular Devices) in the 532 nm channel and the 635 nm channel corresponding to the excitation/emission profiles of Alexa Fluor™ 555 and Alexa Fluor™ 647, respectively. Raw fluorescence signals and background signals were automatically generated by the Genepix Pro™ 7 software (Molecular Devices). Raw data were processed and analyzed with a custom script as previously described.⁸⁸

Enrichment of serum glycoproteins by lectins

All steps in this section (2.4.2) were performed at room temperature in triplicates. Pooled serum was prepared by combining 2 µL of each of 96 randomly chosen serum samples. Total protein concentration of the pooled sample was measured with *DC*™ protein assay (Bio-Rad Laboratories).

For AIA pulldown, 500 µL AIA-agarose beads (~0.5 mg AIA, Vector Laboratories), was added to a spin column. The beads were washed with 400 µL PBS. Pooled serum containing 2 mg of total protein was diluted to 400 µL with PBS and

incubated with the beads for 1 hour with gentle agitation. The beads were washed with 400 μ L PBS three times. To elute glycoproteins, 100 μ L of Glycoprotein Eluting Solution (Galactose/GalNAc) (Vector Laboratories) was added to the column. The beads were incubated with the eluting solution for 15 minutes, and the eluate was collected. This eluting step was repeated, and the two eluate fractions were combined.

For diCBM40 pulldown, Ni-NTA magnetic beads (Thermo Fisher Scientific) were prepared by washing 200 μ L of beads (~2.5 mg) with 400 μ L PBS. 40 μ L of diCBM40 solution (~240 μ g His-tagged diCBM40) were immobilized on the column by incubating for 30 minutes with gentle agitation. Pooled serum containing 1 mg of total protein was diluted to 360 μ L with PBS, mixed with the lectin-conjugated beads, then incubated altogether for 30 minutes with gentle agitation. The beads were washed with 400 μ L wash buffer (20mM sodium phosphate, 150mM NaCl, 25mM imidazole, pH = 8.0) three times. To elute glycoproteins, 100 μ L of elution buffer (20mM sodium phosphate, 150mM NaCl, 400mM imidazole, pH = 8.0) was added to the beads. The beads were incubated with the eluting solution for 15 minutes, and the eluate was collected. This eluting step was repeated, and the two eluate fractions were combined.

Mass spectrometry

(Note: mass spectrometry was done by Jack Moore at the Alberta Proteomics and Mass Spectrometry Facility.)

40 μ L of eluted solutions were electrophoresed by SDS-PAGE until dye front travelled \sim 1cm into the gel. Gel bands containing all proteins were excised and rinsed with deionized water.

In-gel trypsin digestion was performed on the samples. Samples were reduced (10mM BME in 100mM bicarbonate) and alkylated (55mM iodoacetamide in 100mM bicarbonate). After dehydration enough trypsin (6ng/ μ L, Promega Sequencing grade) was added to just cover the gel pieces and the digestion was allowed to proceed overnight (\sim 16 hrs.) at 37C. Tryptic peptides were first extracted from the gel using 97% water/2% acetonitrile/1% formic acid followed by a second extraction using 50% of the first extraction buffer and 50% acetonitrile.

The tryptic peptides were resolved and ionized by using nano flow HPLC (Easy-nLC 1000, Thermo Scientific) coupled to a Q Exactive Orbitrap mass spectrometer (Thermo Scientific) with an EASY-Spray capillary HPLC column (ES800A, 75 μ m x 15cm, 100 \AA , 3 μ m, Thermo Scientific). The mass spectrometer was operated in data-dependent acquisition mode, recording high-accuracy and high-resolution survey orbitrap spectra using external mass calibration, with a resolution of 35,000 and m/z range of 300–1700. The twelve most intense multiply charged ions were sequentially fragmented by using HCD dissociation, and spectra of their fragments were recorded in the orbitrap at a resolution of 17,500; after fragmentation all precursors selected for dissociation were dynamically excluded for 30 s. Data was processed using Proteome Discoverer 1.4 (Thermo Scientific) and the database, Uniprot Human UP000005640, was searched using SEQUEST (Thermo Scientific). Search parameters included a strict

false discovery rate (FDR) of .01, a relaxed FDR of .05, a precursor mass tolerance of 10ppm and a fragment mass tolerance of 0.01Da. Peptides were searched with carbamidomethyl cysteine as a static modification and oxidized methionine and deamidated glutamine and asparagine as dynamic modifications.

Chapter 3. Discovery of Glycan Biomarkers in COVID-19

3.1 Introduction

COVID-19 is the disease caused by SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) infection. The World Health Organization estimates that the COVID-19 pandemic has resulted in > 6,300,000 deaths worldwide as of June 2022.¹²⁸ Severity of COVID-19 can vary significantly among the population, ranging from no symptoms or mild respiratory symptoms to ARDS (acute respiratory distress syndrome) and multi-organ dysfunction.^{129,130} Extensive studies since the pandemic have pointed to a central role of immune system dysregulation in the development of severe COVID-19. Signs of such dysregulation at molecular levels can include excessive release of pro-inflammatory cytokines (the “cytokine storm”), increased complement system activation, accumulation of neutrophil extracellular traps (NETs), and augmented autoantibody reactivities.^{131–135}

As elaborated in Chapter 1, host glycans are essential in inflammatory and immune response in infectious diseases. It is therefore plausible to assume an association between host glycosylation and severe COVID-19, in which the host immune system is dysregulated. Indeed, studies have shown IgG glycosylation is associated with COVID-19 severity. Two independent studies found decreased core fucosylation in SARS-CoV-2-specific IgG in severe COVID-19 cases, compared to mild cases.^{70,71} It was argued that lack of core fucosylation in IgG enhances affinity to the Fc receptor FcγIIIR on myeloid cells, which promotes the hyperinflammatory responses in

severe COVID-19.⁷¹ In addition to core fucosylation, decrease in galactosylation and increase in sialylation of antigen-specific IgG also coorelated with more severe disease.⁷⁰ In another study analyzing glycosylation of total plasma IgG, lower levels of bisecting GlcNAc were associated with severe COVID-19 in one of the three cohorts examined.¹³⁶

It is understandable that all previous studies focused on the glycosylation of IgG given the critical role of antibodies in infectious diseases including COVID-19.¹³⁷ Nevertheless, antibodies are only one of the many aspects of severe COVID-19 pathogenesis. Moreover, IgG response does not synchronize with disease progression. For most patients, IgG seroconversion occurs one to two weeks after symptom onset and IgG titer peaks at week 3.¹³⁷ Whether glycosylation in early disease phase correlates with prognosis remained unknown. Aiming to address these issues, I generated the plasma glycomic profiles of COVID-19 patients stratified by severity using lectin microarray. Glycomic findings in plasma were corroborated by glycomic data from post-mortem tissues of COVID-19 patients and followed up by histological and glycoproteomic analyses (work by collaborators). In brief, increased α 2,6-sialylation, along with by increased ST6GAL1, is the most notable signature in severe COVID-19 cases compared to mild cases and negative controls. α 2,6-sialylation was linked to extensive complement protein deposition and related pathological features in the tissues from COVID-19 autopsies. Overall, this work identified α 2,6-sialic acid as a prognosis marker in the early phase of the disease and suggests a potential role of α 2,6-sialylation in complement system dysfunction in severe COVID-19.

3.2 Results and Discussion

3.2.1 α 2,6-sialylation is increased in severe COVID-19 patient plasma

Plasma glycomes from a COVID-19 patient cohort and a sex/age-matched negative control cohort (**Table 3.1**) were collected. The COVID-19 patient cohort was comprised of 71 individuals who were PCR-tested positive for SARS-CoV-2, and stratified into three groups based on endpoint disease severity: (i) patients who were not hospitalized (*mild*, n = 5), (ii) patients who were hospitalized but did not need supplemental oxygen (*moderate*, n = 8) and (iii) patients who were hospitalized, received supplemental oxygen and/or were admitted into the intensive care unit (ICU) (*severe*, n = 58). All COVID-19 plasma samples were collected at the University of Alberta Hospital (Edmonton, Alberta, Canada) and at their first visit to the hospital. COVID-19 negative control plasma samples were from a healthy cohort (n = 60) recruited at the University of Georgia (Athens, Georgia, USA).

Table 3.1. Descriptive characteristics of the plasma cohorts.

	COVID-19 Positive (n=71)	COVID-19 Negative (n=60)
Male/Female (ratio)	39/30 (1.30)	35/25 (1.40)
Median Age in Years (IQR)	61 (50-73)	61 (50-71)
Collection Date Ranges: YYYY/MM-YYYY/MM	2020/07-2020/09 (n=2) 2020/10-2021/01 (n=64) 2021/02-2021/04 (n=5)	2019/09-2019/10 (n=60)

Note: The gender information of 2 participants and age information of 2 participants the COVID-19 positive cohort are not available.

To investigate the glycomic differences between different categories of patients, plasma samples were analyzed by lectin microarray (list of microarray lectins available in the appendices) (**Figure 3.1a**). A heatmap overview of the glycomic data of the cohort is presented in **Figure 3.1b**. Comparison of lectin microarray data between patient groups revealed significant differences in plasma sialic acid. Plasma from severe COVID-19 patients was significantly higher in α 2,6-sialic acid with respect to the negative controls and the mild COVID-19 patients (SNA, **Figure 3.1c**). SNA binding to severe COVID-19 plasma was almost twice (~84% increase) that of the mild. In recent work on adult ferret model of influenza infection, increased α 2,6-sialylation was observed at early infections, suggesting some similarities between the glycomic responses to SARS-CoV-2 and influenza infection.⁸⁹ In contrast to α 2,6-sialylation, α 2,3-sialylation was significantly lower in severe COVID-19 patients with respect to the negative controls, as indicated by SLBR-H binding in **Figure 3.1c**. Additionally, SLBR-B, which prefers a narrower subset of *O*-linked α 2,3-sialic acid,¹⁰³ also showed decreased binding to severe COVID-19 patients (**Figure 3.2a**). Data from the α 2,3-sialic acid-binding probe diCBM40 did not exhibit differences between sample groups (**Figure 3.2a**), likely due to the offsetting effect of α 2,6-sialic acid which can bind diCBM40 with lower affinity.¹⁰⁸ Besides sialylation, differences in plasma *O*-glycans were observed. AIA and MPL, both recognizing core 1/3 *O*-glycans, had lower binding to severe COVID-19 plasma compared to the negative controls (**Figure 3.2b**), matching the observation in SLBR-B.

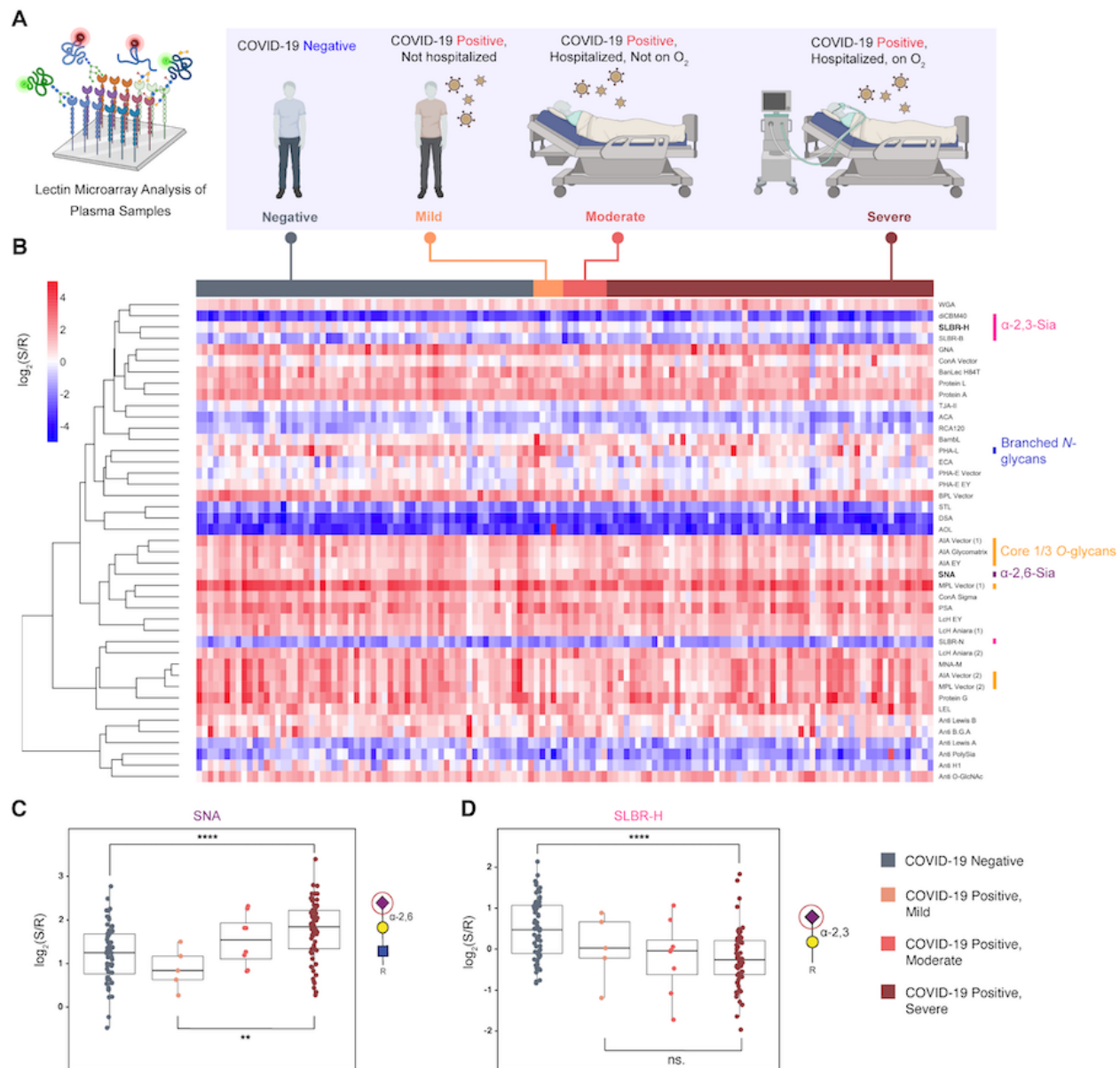


Figure 3.1. Plasma glycomic profiles of COVID-19 positive and negative cohorts. **(a)** Schematic description of analysis. COVID-19 positive patients were categorized into three groups by disease severity (Mild, Moderate, Severe) and compared to an age and gender-matched control cohort (Negative). Plasma samples were analyzed by lectin microarray. **(b)** Heatmap of lectin microarray data with annotations of rough glycan specificities for select lectins. Columns (patients) are ordered by disease severity as in (a), indicated at the top of the heatmap. **(c)** Boxplot analysis of SNA and SLBR-H binding data by patient group. Mann-Whitney U test was used to determine p values. **: $p < 0.01$; ****: $p < 0.0001$; ns: $p > 0.05$.

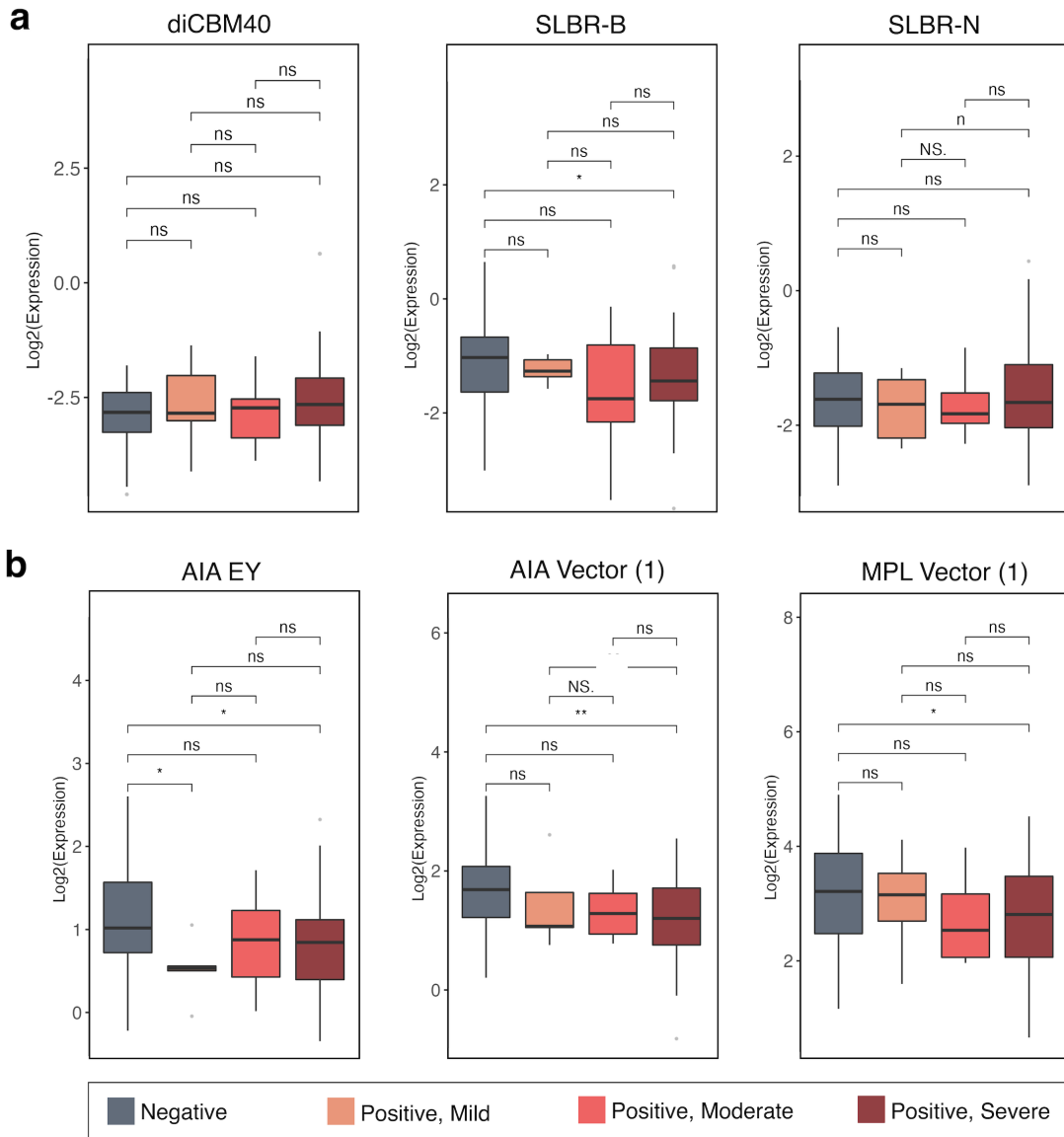


Figure 3.2. Boxplot analysis of select **(a)** α 2,3-sialic acid-binding lectins and **(b)** core 1/3 O-glycan-binding lectins between patient groups. Legend at the bottom of the figure applies to all panels. “Negative” and “Positive” denote COVID-19 status. “Mild”, “Moderate” and “Severe” denote COVID-19 severity. Mann-Whitney U test was used to determine p values. *: $p < 0.05$; **: $p < 0.01$; ns: $p \geq 0.05$.

3.2.2 α 2,6-sialylation and its biosynthesis are increased in post-mortem COVID-19 patient tissues

(Note: experiments described in 3.2.2 were performed by collaborators including Dr. Emma Kurz and Dr. Shuhui Chen)

Severe COVID-19 is a multi-organ disease. It can be complicated with coagulation abnormalities, kidney injury, heart disease, etc.^{138–141} Inflammation and myeloid cell infiltration have been observed in affected tissues including the lung, indicative of immune system dysregulation. To investigate whether glycosylation changes can also be observed in the tissues of severe COVID-19 patients, tissue (heart, kidney, liver, upper lobe lung, lower lobe lung) glycomes of post-mortem COVID-19 autopsy tissues and negative control tissues were generated with lectin microarray. The COVID-19 positive cohort was comprised of patients died of from COVID-19 in New York City, USA (**Table 3.2**). The control cohort was comprised of COVID-19 negative individuals who deceased due to other causes. Of note, most of the COVID-19 negative cohort decedents exhibited pulmonary pathology at the time of decease (**Table 3.2**). Autopsy tissue quality was assured by the integrity of nuclei in hematoxylin and eosin staining shown in **Figure 3.3a**.

Table 3.2 Descriptive characteristics of the patients of the autopsy tissue cohorts.

Patient ID	Age	Sex (M/F)	COVID-19 Status	Diagnoses	Pulmonary Pathology (Y/N)
A14-00019	57	M	Negative	-DAD; pleural effusion -MI; hypertrophy -Cirrhosis, Hep C+	Y
WA19-00062	64	M	Negative	-Subarachnoid Hemorrhage, internal carotid artery dissection	N
A19-00028	9	F	Negative	-GBM -Aspiration pneumonia (Staph Aureus) -Cardiomegaly	Y
A13-00016	78	F	Negative	-DAD -MI; coronary artery disease	Y
A18-00057	76	F	Negative	-CHF -intra-abdominal hemorrhage -CKD; pneumonia	Y
A08-00032	79	F	Negative	-DAD; pulmonary infarcts, lung congestion**	Y
A19-00047	47	M	Negative	-Cor pulmonale due to pericarditis and hemorrhage; -pulmonary emboli	Y
A19-00051	67	M	Negative	-pleural fluid -Coronary artery disease -hepatomegaly	N
WA19-00072	75	M	Negative	-Disseminated intravascular coagulation to sepsis -DAD; cardiomegaly	Y
A19-00048	76	M	Negative	-Hypovolemic shock -CHF -pulmonary edema with pleural effusions	Y
20-12	64	F	Positive	-COPD, lung cancer	Y
20-20 (18)	56	F	Positive	N/A	Y

20-21	71	F	Positive	-Coronary artery disease, adrenal adenoma -COPD	Y
20-17	64	F	Positive	-Renal cell carcinoma, partial nephrectomy -DAD	Y
20-13	60	M	Positive	-CAD -acute MI	Y
20-15	50	F	Positive	-Hyperthyroidism -DAD	Y
20-16	44	M	Positive	-DAD -Acute Kidney Injury	Y
20-19	65	M	Positive	-Cirrhosis, CKD -PVD	Y

Key: DAD = Diffuse Alveolar Damage; MI = myocardial ischemia; CHF = congestive heart failure; CKD = chronic kidney disease; CAD = coronary artery disease

**Autopsy Limited to Lungs Only.

In line with the data from plasma samples, higher levels of α 2,6-sialic acids were observed in the lower lobe lungs of COVID-19 patients (SNA, **Figure 3.3b**). This pattern, however, was not observed in the upper lobe lungs. This was in concordance with radiographic findings showing viral pneumonia including SARS-CoV-2 pneumonia mainly affects the lower lobe lung.^{142–144} Opposite to the data from plasma samples, levels of α 2,3-sialic acids also seemed to increase in COVID-19 patient tissues including the upper lobe lungs, the lower lobe lungs and livers (diCBM40, **Figure 3.3b**). No differences in glycosylation were observed in the heart and kidney tissues samples (**Figure 3.3b**). These results were then confirmed by lectin fluorescence staining using the same lectins and tissue samples. As shown in **Figure 3.4a** and **Figure 3.4b**,

increased α 2,6-sialylation was observed in both the upper and lower lobe of the lungs of the COVID-19 patients. The magnitude of increase in upper lobe lung was less than the lower lobe. Increased α 2,3-sialylation was observed in both lobes of the lungs of the COVID-19 patients (**Figure 3.4c**, **Figure 3.4d**).

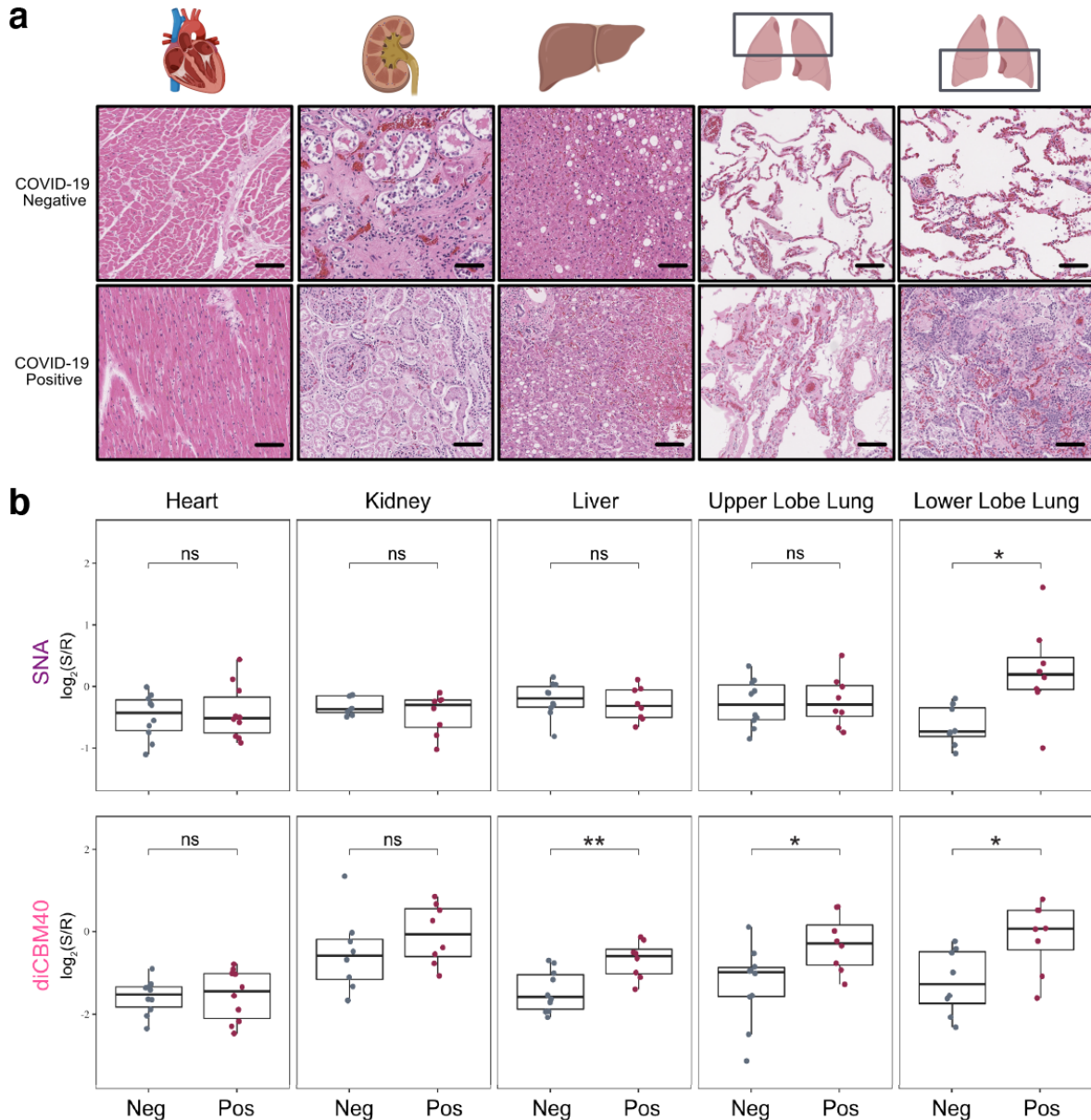


Figure 3.3. Glycosylation patterns in tissues (heart, kidney, liver, upper and lower lobe lung) of deceased COVID-19 patients and negative controls. **(a)** Representative images of hematoxylin and eosin stains of tissues from COVID-19 positive and COVID negative autopsy samples. Scale bars represent 75 μ m. **(b)**

Boxplot analysis of SNA and diCBM40 binding data by patient group. Student's t-test was used to determine p -values. *: $p < 0.05$; **: $p < 0.01$; ns: $p \geq 0.05$. Modified from the work of Dr. Shuhui Chen and Dr. Emma Kurz.

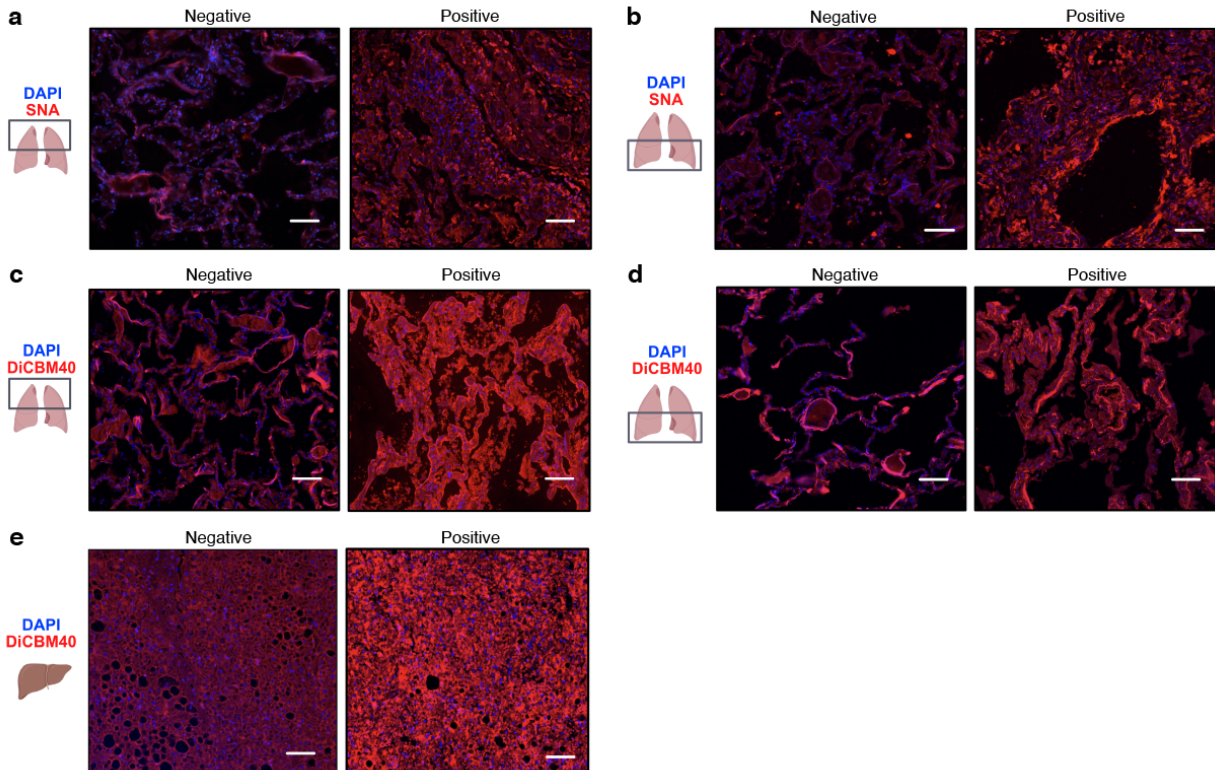


Figure 3.4. Representative images of lectin fluorescence staining (red) and nuclei staining (blue) in tissues of deceased COVID-19 patients and negative controls. **(a)** SNA staining in upper lobe lungs; **(b)** SNA staining in lower lobe lungs; **(c)** diCBM40 staining in upper lobe lungs; **(d)** diCBM40 staining in lower lobe lungs; **(e)** diCBM40 staining in livers. All scale bars represent 150 μm . Modified from the work Dr. Emma Kurz.

My work and work by collaborators both showed host $\alpha 2,6$ -sialylation increase may be a marker of severe COVID-19. Increase in this type of glycosylation can be because of increased protein synthesis, increased activity of $\alpha 2,6$ -sialylation machinery, or both. To obtain clues of the origin of this $\alpha 2,6$ -sialylation increase, the lung and liver tissues were stained for beta-galactoside alpha-2,6-sialyltransferase 1 (ST6GAL1), the

main enzyme that biosynthesizes the α 2,6-sialic acid linkage. Expression ST6GAL1 was significantly higher in the lungs of COVID-19 patients, especially in the epithelium (**Figure 3.5a, Figure 3.5b**). ST6GAL1 staining was extensive in the liver, matching previous studies showing the liver expresses the highest amount of ST6GAL1 among all tissues.^{145–147} However, no significant differences were observed in the liver between the two categories of patient samples (**Figure 3.5c**). Other studies also found evidence of ST6GAL1 upregulation in the liver, lung epithelium, and immune cells in COVID-19 patients.^{148–150} In addition to increased activity of α 2,6-sialyltransferases, upregulation of α 2,6-sialic acid could also be due to altered activity of host neuraminidases, which could be investigated in future studies.

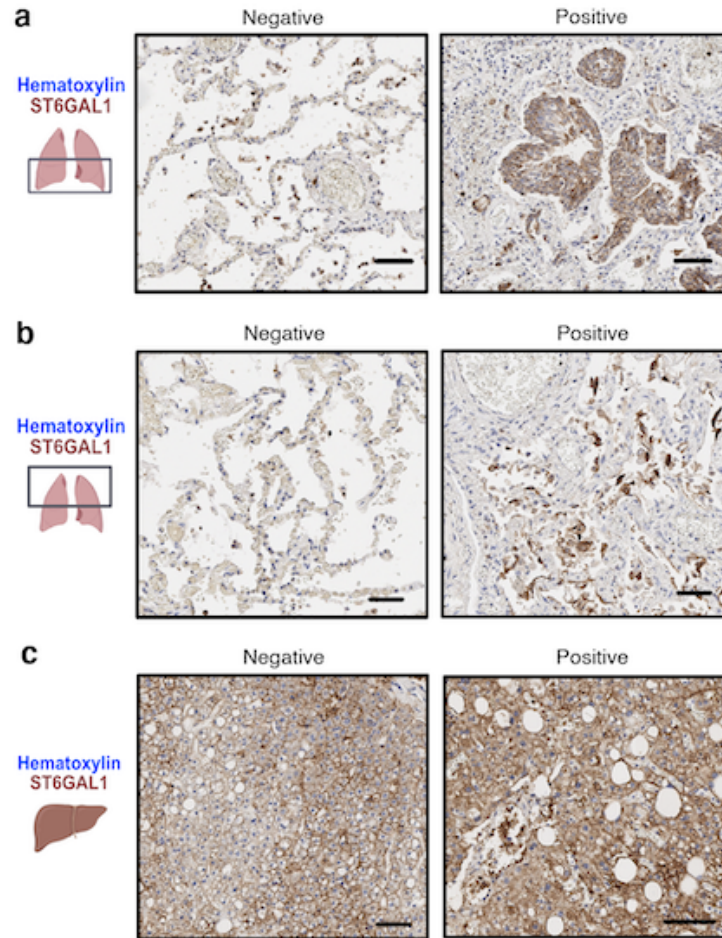


Figure 3.5. Representative images of ST6GAL1 staining (brown) and hematoxylin staining (blue) in the **(a)** upper lobe lungs, **(b)** lower lobe lungs and **(c)** livers of deceased COVID-19 patients and negative controls. All scale bars represent 150 μ m. Modified from the work Dr. Emma Kurz.

3.2.3 Differential α 2,6-sialylation of complement proteins is associated with COVID-19 severity

(Note: IHC staining experiments described in 3.2.3 were performed by collaborators including Dr. Emma Kurz and Dr. Shuhui Chen)

Glycomic analyses above identified an association between increased α 2,6-sialic acid and severe COVID-19. To gain insights into the potential mechanisms behind this association, pooled plasma samples corresponding to the severe COVID-19 and mild COVID-19 category were prepared and enriched with SNA. Compositions of the enriched glycoproteins were resolved by mass spectrometry (**Figure 3.6a**). There were 77 plasma proteins enriched by SNA in the severe COVID-19 sample and 38 enriched in the mild COVID-19 sample (**Table 3.3**). This is consistent with the higher α 2,6-sialylation level in the severe disease category. Shared between the two enriched samples were 29 proteins. In the severe COVID-19 sample, 44 proteins enriched by SNA were more abundant than in the mild sample, the top 25 (by abundance in the severe COVID-19 sample) of which are shown in **Figure 3.6b**. Of note, proteins in the blood coagulation pathway, such as fibrinogen, plasminogen, and prothrombin, were significantly higher in the severe COVID-19 sample. This could be due to an increase in the expression of these proteins, which has been identified in multiple studies and was believed to contribute to the coagulopathy observed in severe cases of COVID-19.^{151–}

154

The next group of SNA-enriched and most upregulated proteins in severe COVID-19 plasma are complement proteins, including several downstream proteins in the complement cascade (e.g., C5, C6, C7, C8, C9) and some upstream ones (e.g., C1r). In fact, Gene Ontology pathway enrichment analysis using the differentially enriched proteins as input showed complement-related pathways were the most significantly enriched (**Figure 3.6c**). Immunoglobulins, which are abundant and α 2,6-

sialylated,⁸ did not show different enrichment by SNA between samples of different COVID-19 severity. This may be because the plasma samples were collected before significant seroconversion.

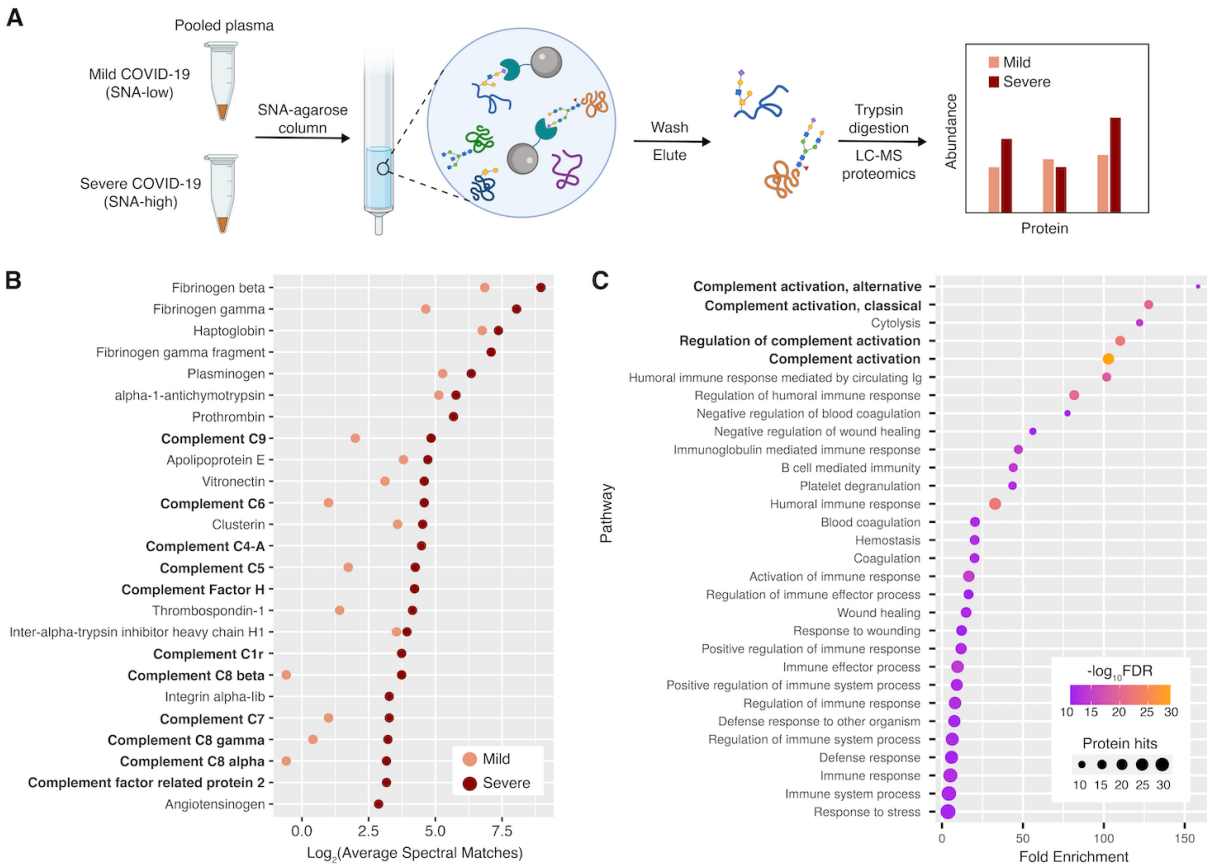


Figure 3.6. Glycoproteomic analysis of α 2,6-sialic acid-containing proteins from mild and severe COVID-19 plasma. **(a)** Scheme of workflow. Mild COVID-19: COVID-19 patients who were not hospitalized. Severe COVID-19: COVID-19 patients who were hospitalized and received supplemental oxygen. **(b)** SNA-reactive glycoproteins significantly enriched the severe COVID-19 plasma compared to mild, and their mass spectrometric abundance profiles (average spectral matches). The top 25 enriched glycoproteins (by abundance) in the severe group, are shown. **(c)** Pathway enrichment analysis of the enriched plasma glycoproteins. FDR: false discovery rate.

Table 3.3 List of SNA-enriched proteins in the pooled severe and mild COVID-19 plasma

Protein Name	Uniprot Accession	Average Number of Spectral Matches (Severe)	Average Number of Spectral Matches (Mild)
Fibrinogen beta chain	P02675	496	0
Fibrinogen alpha chain	P02671	292	0
Fibrinogen gamma chain	P02679	264	25
Complement C3	P01024	236	189
Apolipoprotein B-100	P04114	193	171
Alpha-1-antitrypsin	P01009	157	137
Alpha-2-macroglobulin	P01023	141	243
Fibrinogen gamma chain (Fragment)	C9JPQ9	136	0
Complement factor H	P08603	125	104
Plasminogen	P00747	81	39
Hemopexin	P02790	73	97
Immunoglobulin kappa constant	P01834	67	0
Alpha-1-antichymotrypsin	P01011	55	35
Prothrombin	P00734	51	0
Immunoglobulin heavy constant mu	P01871	45	60
Inter-alpha-trypsin inhibitor heavy chain H4	Q14624	45	43
Apolipoprotein A-IV	P06727	39	0
Immunoglobulin heavy constant alpha 2 (Fragment)	A0A0G2JMB2	36	0
C3/C5 convertase	B4E1Z4	36	0
Fibronectin	P02751	36	0
Complement component C9	P02748	29	0
Apolipoprotein E	P02649	26	14
Vitronectin	P04004	24	0
Complement component C6	P13671	24	0
Clusterin	P10909	23	0
Alpha-1B-glycoprotein	P04217	23	0
Complement C4-A	P0C0L4	22	0
Leucine-rich alpha-2-glycoprotein	P02750	20	0
Complement C5	P01031	19	0
Complement factor H	A0A0D9SG88	19	0
Gelsolin	P06396	18	0
Thrombospondin-1	P07996	18	0

Lipopolysaccharide-binding protein	P18428	17	11
Hemoglobin subunit alpha	P69905	17	17
Alpha-2-HS-glycoprotein	P02765	16	27
Centrosomal protein of 290 kDa (Fragment)	A0A5K1VW81	16	0
Inter-alpha-trypsin inhibitor heavy chain H1	P19827	15	12
Complement subcomponent C1r	A0A3B3ISR2	13	0
Complement component 8 subunit beta	F5GY80	13	0
Alpha-2-antiplasmin	P08697	13	10
Immunoglobulin lambda constant 2	P0DOY2	13	0
Complement factor I	G3XAM2	12	10
Pregnancy zone protein	P20742	12	16
Serum amyloid P-component	P02743	11	0
Integrin alpha-IIb	P08514	10	0
Complement component C7	P10643	10	0
Complement component C8 gamma chain	P07360	9	0
Uncharacterized protein C9orf43	Q8TAL5	9	0
Platelet factor 4	P02776	9	0
Complement component C8 alpha chain	P07357	9	0
Complement factor H-related protein 2	V9GYE7	9	0
Fermitin family homolog 3	Q86UX7	9	0
Afamin	P43652	8	23
Angiotensinogen	A0A7P0T9S6	7	0
C-reactive protein	P02741	6	0
Carboxypeptidase N catalytic chain	P15169	6	0
Ficolin-3	O75636	5	0
Pigment epithelium-derived factor	P36955	5	0
Plasma kallikrein	P03952	5	7
Pleckstrin	P08567	5	0
Apolipoprotein L1	O14791	4	4
Plasma serine protease inhibitor	P05154	4	3
Thyroxine-binding globulin	P05543	4	5
Coagulation factor V	P12259	3	0

2-phospho-D-glycerate hydro-lyase	A0A2R8Y6G6	3	0
Potassium voltage-gated channel subfamily H member 1	A0A0S1TJ81	3	0
N-acetylmuramoyl-L-alanine amidase	Q96PD5	3	0
Immunoglobulin lambda-like polypeptide 5	B9A064	3	5
Integrin beta (Fragment)	H3BM21	3	0
Apolipoprotein C-II	P02655	3	4
Centromere-associated protein E	A0A087X0P0	2	0
Corticosteroid-binding globulin (Fragment)	G3V4V7	2	0
Beta-2-microglobulin	P61769	2	2
Ig-like domain-containing protein (Fragment)	A0A0J9YY99	2	5
Protein ITPRID2	E7EUL7	2	3
Protein mono-ADP-ribosyltransferase PARP6 (Fragment)	H3BUQ6	2	0
Insulin-like growth factor-binding protein complex acid labile subunit	P35858	2	0
Haptoglobin	P00738	0	108
Haptoglobin-related protein	P00739	0	70
Kininogen-1	P01042	0	53
Antithrombin-III	P01008	0	49
Serum amyloid A-1 protein	P0DJI8	0	2
Complement factor H-related protein 1	Q03591	0	19
SAA2-SAA4 readthrough	A0A096LPE2	0	8
Heparin cofactor 2	P05546	0	10
Serum paraoxonase/arylesterase 1	P27169	0	5

The pathway enrichment analysis result suggested that α 2,6-sialic acid upregulation may be mechanistically linked to severe COVID-19 via the complement system. Complement activation in COVID-19 has been reported in multiple studies and correlates with severity.^{135,155,156} The complement cascade promotes inflammation and generates pro-inflammatory cytokines, which lead to tissue damage.⁴² Accumulation of

neutrophil extracellular traps (NETs) and increased myeloid cell-mediated inflammation observed in severe COVID-19 can also be attributed to overstimulation of the complement system.¹⁵⁷ Targeting the complement system has been receiving increasing attention as a potential route to severe COVID-19 management.^{158–160}

Considering the previous findings on augmented complement activation in severe COVID-19 and the upregulation of α 2,6-sialylation in severe COVID-19 patients, I next investigated whether the complement proteins were more sialylated in the severe COVID-19 plasma. C5 and C9 were selected for this purpose because they were more enriched in the severe COVID-19 plasma, and their derivatives are commonly probed to measure complement activity in tissues. SNA enrichment was performed on the same pooled plasma for glycoproteomic analysis and a pooled plasma of the negative controls, then C5 and C9 were detected by Western blot. Compared to the negative control, COVID-19 plasma, regardless of severity, showed higher C5 and C9 α 2,6-sialylation relative to the total protein levels (**Figure 3.7**). This showed the α 2,6-sialylated fractions of plasma C5 and C9 were increased in COVID-19 compared to non-disease state. Severe COVID-19 patient plasma showed higher α 2,6-sialylation of C5 compared to the mild, which may be indicative of severity-dependent difference in α 2,6-sialylation (**Figure 3.7**).

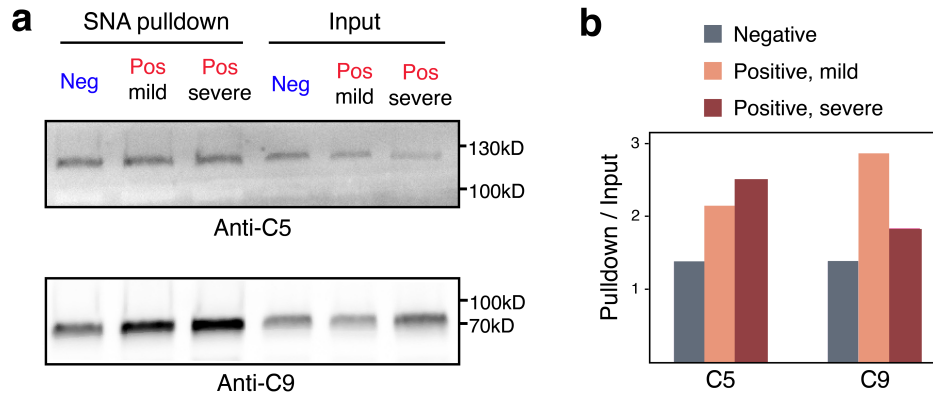


Figure 3.7. Differential α 2,6-sialylation of complement C5 and C9 in three patient groups. **(a)** Western blot (anti-C5 and anti-C9) of SNA pulldown samples from pooled patient plasma and corresponding input. **(b)** Bar plot of the pulldown/input band intensity ratios.

Since upregulation of α 2,6-sialylation was observed in the COVID-19 autopsy tissues, whether complement activation is associated with this glycosylation signature was investigated by immunohistochemical (IHC) staining for C5b and C9. Significantly greater C5b and C9 depositions were observed in the lungs and livers of COVID-19 patients compared to the negative controls (**Figure 3.8**). In the lower lobe lungs, most complement deposits were located at airway barriers, coinciding with the findings in SNA fluorescence staining (**Figure 3.4b**). This suggested at least some of the increase in α 2,6-sialic acids came from increased complement protein deposition, as C5b and C9 are sialylated. In addition, complement-mediated inflammation and myeloid cell recruitment in the COVID-19 tissue samples were confirmed by greater IL-6 (interleukin 6) and CD163⁺ macrophage staining (**Figure 3.9**). IL-6 is an established marker of COVID-19 severity, which can both promote and be promoted by complement activation.^{161–163} Complement system activation can lead to recruitment of myeloid cells

(e.g., macrophages),^{164–169} and exacerbated pulmonary fibrosis mediated by CD163⁺ macrophages has been found in severe COVID-19.¹⁷⁰

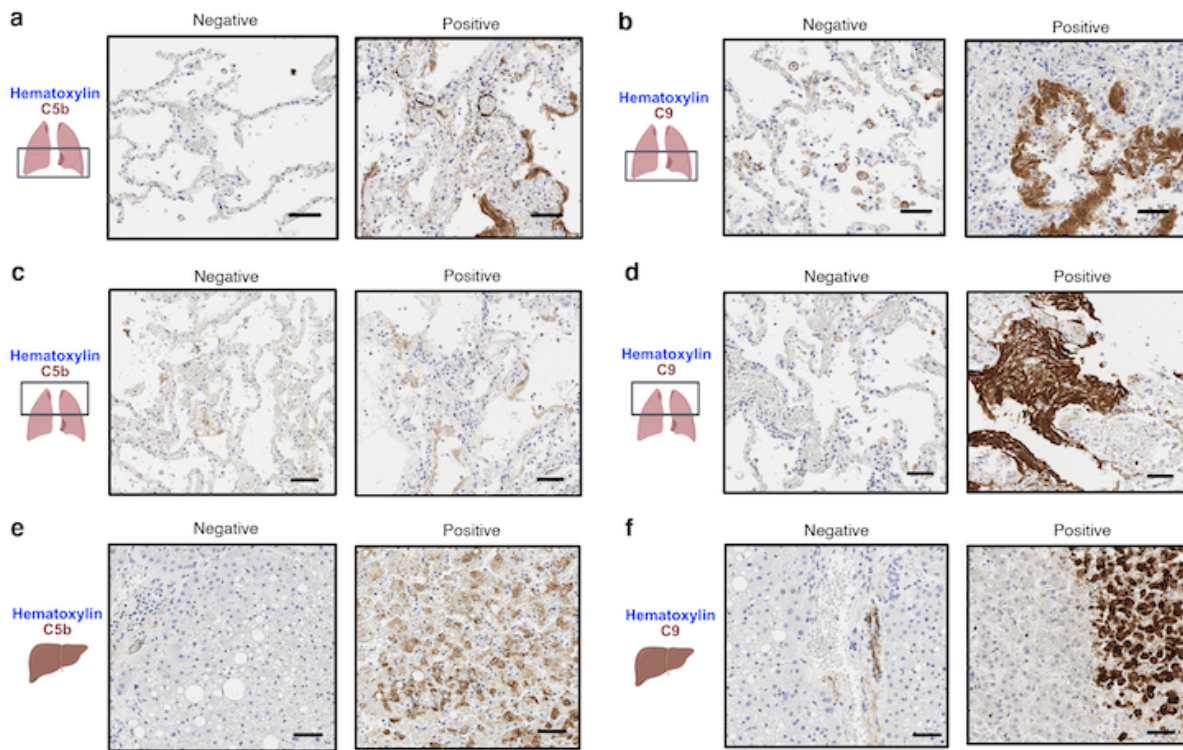


Figure 3.8. Representative images of C5b and C9 staining (red) and nuclei staining (blue) in **(a, b)** lower lobe lungs, **(c, d)** upper lobe lungs and **(e, f)** livers of deceased COVID-19 patients and negative controls. All scale bars represent 150 μm . Modified from the work Dr. Emma Kurz.

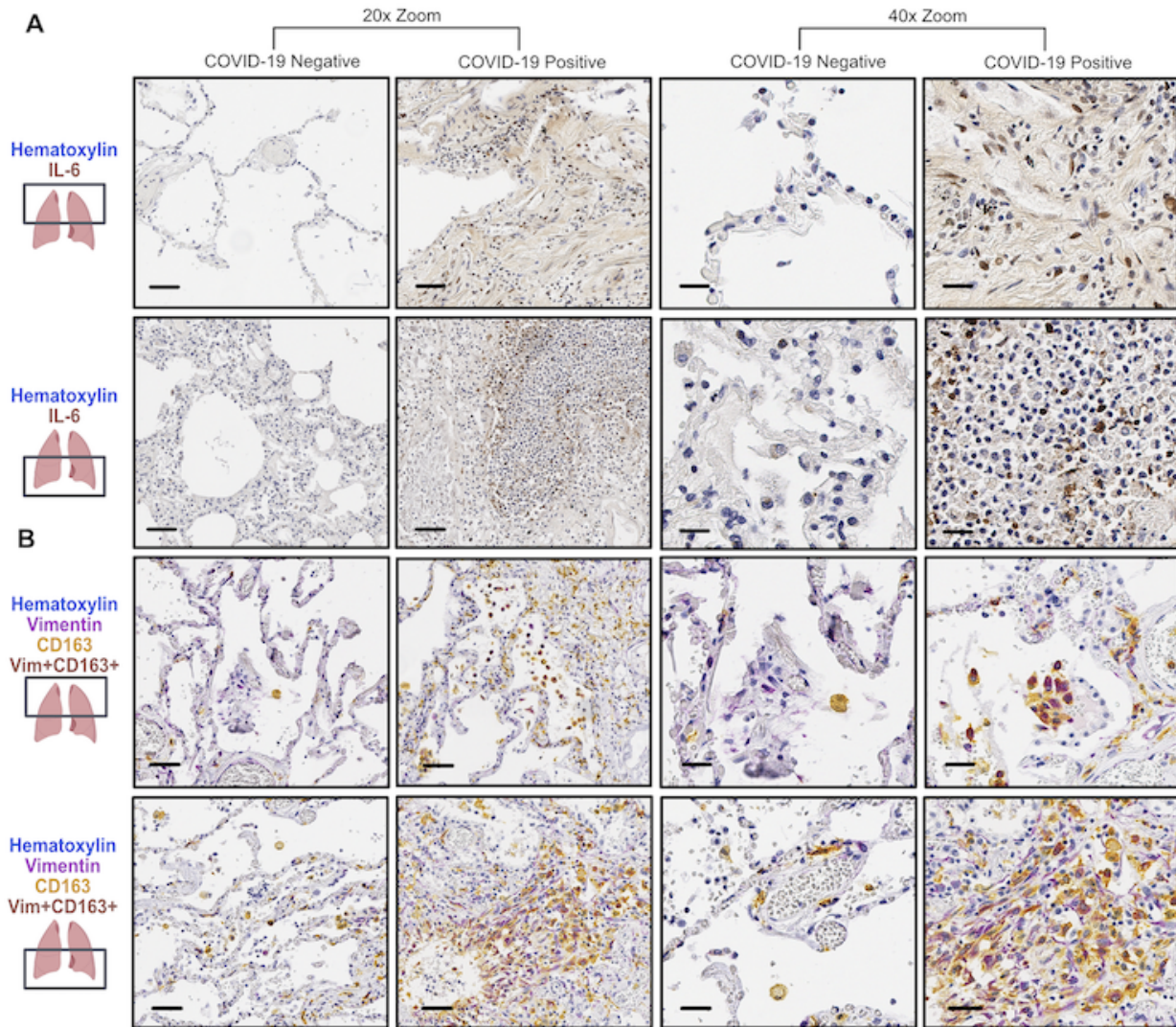


Figure 3.9. Representative images of **(a)** IL-6 staining and **(b)** CD163 (yellow), vimentin (purple) and nuclei (blue) staining in the tissues of deceased COVID-19 patients and negative controls. 20x zoom scale bars represent 100 μ m. 40x zoom scale bars represent 50 μ m. Work Dr. Emma Kurz.

Complement overstimulation and related pathological changes in severe COVID-19 patient tissues were highly relevant to the plasma, as circulating complement proteins directly deposit on tissues when the cascade is activated. Nonetheless, the lowest C5 level was in the severe COVID-19 plasma (**Figure 3.7a**). This agreed with previous reports showing concentrations of plasma complement proteins, including C5

and C9, are not necessarily greater, and even lower sometimes, in severe COVID-19 patients when compared to mild cases or negative controls.^{148,171–174} Therefore, upregulation of complement system activity in COVID-19 cannot be explained by higher expression of complement proteins in plasma. Instead, this study identified an association of higher α 2,6-sialic acid content of C5 and C9 with the hyperactivated complement system in severe COVID-19. Content of α 2,6-sialylation may change the biology of C5 and C9, influencing the potential of complement system activation. Some earlier studies indeed showed that glycosylation state influences the biological activity of C5a and C9. Removal of *N*-glycans from C5a-desArg, a pro-inflammatory anaphylatoxin derived from C5, was significantly more potent.³¹ Whereas de-*N*-glycosylation of C9 resulted in the loss of complement-mediated cell lysis.³³ Functional significance of more specific motifs (e.g., α 2,6-sialic acid) in C5 and C9 glycans has not been explored. Apart from direct functional tuning, α 2,6-sialylation of complement proteins may also increase complement activity by extending their half-lives in the blood, as in the case of IgG.⁴⁸ Another role of glycosylation is altering resistance to protease cleavage.^{92,175} Since the complement cascade is mainly a proteolytic cascade, differential α 2,6-sialylation of complement components may also alter the efficiency of certain proteolytic reactions within the cascade.

3.2.4 α 2,6-sialylation is induced by SARS-CoV-2 infection in ferret lungs

(Note: Ferret tissue collection, viral load determination and RNA-seq experiments described in 3.2.4 were performed by collaborators including Magen Francis and Dr. Alyson Kelvin)

In the patient cohorts, higher α 2,6-sialylation in COVID-19 patients was observed in both the early disease phase (plasma cohort) and more advanced disease phase (tissue cohort). However, it is unclear whether the disparity in α 2,6-sialylation is a pre-existing feature or induced by SARS-CoV-2 infection, which have different clinical implications. To gain some insights into this question, I analyzed the glycome of longitudinal ferret lung tissue samples with our lectin microarray. Ferret is a useful model for influenza studies.¹⁷⁶ In COVID-19 research, this model has been used to study viral transmission, disease symptoms and vaccine response.^{177,178}

Ferrets were infected with SARS-CoV-2 via nasal inoculation of viruses and displayed mild symptoms post-infection. Lung tissues for glycomic analysis were collected on the day of infection (d0), 2 days (d2), 7 days (d7) and 14 days (d14) after infection. Lectin microarray data revealed widespread changes in ferret lung glycosylation post-infection (**Figure 3.10**). Consistent with the findings in patient cohorts, upregulation of α 2,6-sialylation and downregulation of *O*-linked glycans were among the most remarkable glycosylation changes. Of note, α 2,6-sialylation showed an increasing trend with time post-infection (**Figure 3.11a**). This clearly indicated that SARS-CoV-2 infection induced α 2,6-sialylation in ferret lungs.

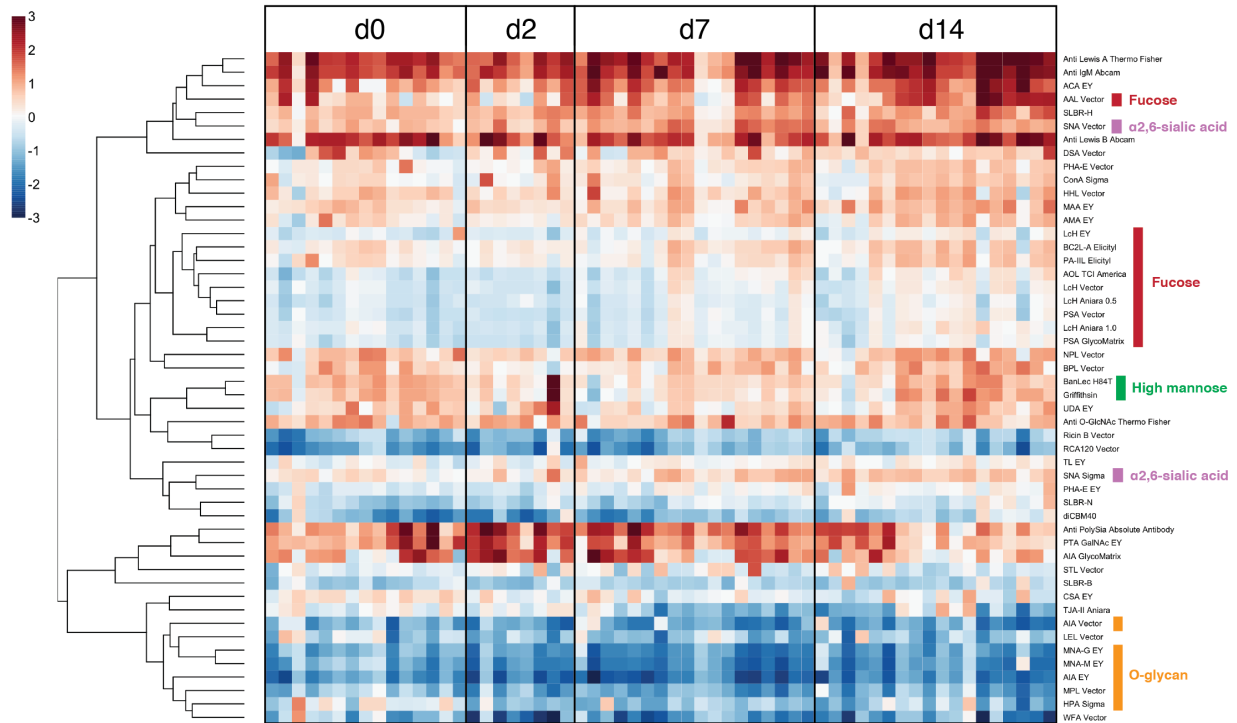


Figure 3.10. Heatmap of lung tissue glycomic profiles of SARS-CoV-2-infected ferrets and negative cohorts. Rough glycan specificities for select lectins are annotated. Columns (patients) are ordered by animals and days post-infection, all indicated in the heatmap.

Gene expression in ferret lung tissues was analyzed by bulk RNA-seq.¹⁷⁹ Transcripts of ST6GAL1 did not change throughout the course of study (**Figure 3.11b**), possibly indicating a mismatch between ST6GAL1 transcript and protein levels (possibly mediated by post-transcriptional regulation) or other regulatory mechanisms of α 2,6-sialylation (e.g., host neuraminidase activity) in the lungs of this COVID-19 model. In the lungs, interferon gamma-mediated antiviral response was strongly activated on d2 and quickly attenuated afterwards (**Figure 3.11b**). As expected, inflammatory response ensued, although only some pro-inflammatory cytokine markers were upregulated (**Figure 3.11b**). Transcriptions of IL-6 and IL-8 significantly elevated after d2,

resembling the temporal change in $\alpha 2,6$ -sialylation. This suggests that $\alpha 2,6$ -sialylation upregulation is accompanied by activation of inflammatory response.

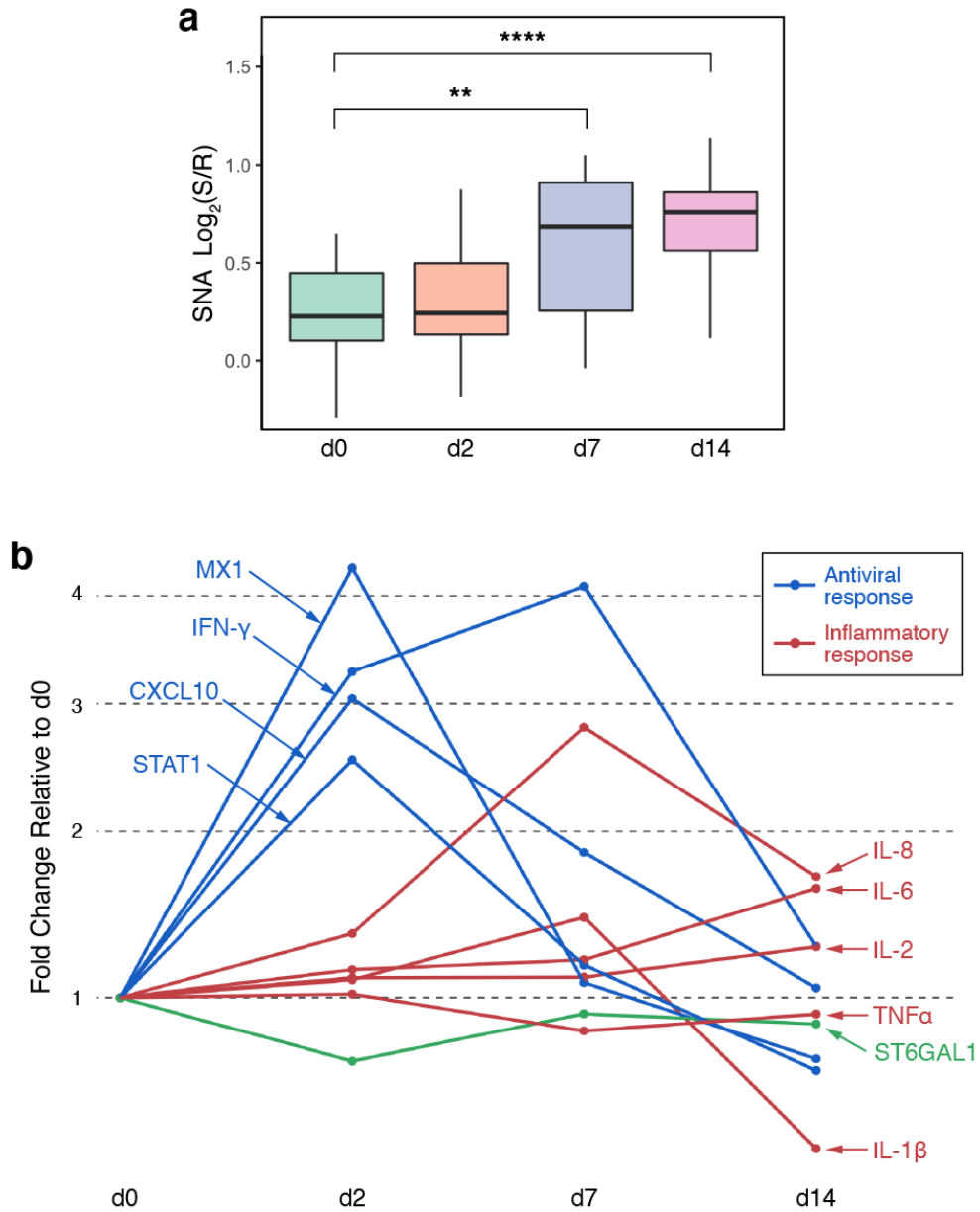


Figure 3.11. Time course changes in $\alpha 2,6$ -sialylation and immune response marker expression in ferret lung tissues. **(a)** Changes in $\alpha 2,6$ -sialylation as indicated by SNA binding. Student's t-test was used to determine p -values. **: $p < 0.01$; ****: $p < 0.0001$. **(b)** Changes in average expression of ST6GAL1 (green), antiviral response markers (blue) and inflammatory response markers (red) from bulk RNA-seq.

3.3 Conclusions

Identification of mechanisms driving hyperinflammation and immune system dysregulation is central to disease management and therapeutic development in severe COVID-19. Plenty of studies since the onset of the pandemic have found a series of severity markers (e.g., IL-6, D-dimer) closely associated with inflammatory response and immune function, which guided the development and clinical trials of therapeutics targeting related biological pathways.^{163,180} In this study, protein α 2,6-sialylation, in both plasma and tissues, was identified as a novel glycan marker of severe COVID-19. Glycoproteomic and histological analysis showed increased α 2,6-sialylation of C5 and C9 is associated with overactivation of the complement system, a typical feature in COVID-19. This study points to the underappreciated roles of glycosylation of complement proteins. As discussed in Chapter 1, while many complement proteins bear structurally diverse glycans, any functional impacts glycosylation may have on the complement proteins and the overall complement system activity in health and diseases have rarely been examined. Although whether and how the increased α 2,6-sialylation of C5 and C9 contributes to overstimulation of complement system in COVID-19 will require more studies to elucidate, this work has arguably shown the potential of integrated glycomic-glycoproteomic analysis in facilitating the discovery of glycan-mediated mechanisms in health and diseases. In addition, glycomic analysis of COVID-19 ferret model showed the SARS-CoV-2 infection elevated α 2,6-sialylation in the lungs. This does not exclude the possibility that high α 2,6-sialylation is a pre-existing

feature in severe COVID-19 patients, as α 2,6-sialylation of serum proteins can be a signature of obesity (see Chapter 2), a risk factor of severe COVID-19.¹⁸¹ Pre-existing high α 2,6-sialylation and infection-induced α 2,6-sialylation may have both contributed to the poor prognosis in some COVID-19 patients.

3.4 Methods

Note: Most of this section (3.4) are the methods of the experiments performed by myself. Methods of the experiments by collaborators (i.e., IHC staining) can be found in corresponding references: ^{179,182}.

Cohorts and Sample Collection

COVID-19 plasma samples were collected from 71 patients recruited from the Intensive Care Unit, the hospital ward or the outpatient clinic at the University Hospital (Edmonton, Alberta, Canada). The CoCollab study was reviewed and approved by the Research Ethics Board/Alberta Research Information Services (ARISE) at the University of Alberta. Recruits were informed of the details of the study by the study team, had the opportunity to ask questions, then signed informed consent. Plasma samples analyzed in this study were collected at the time of enrollment. Blood samples were processed within one hour where possible to isolate plasma, and peripheral blood mononuclear cells, then aliquoted into 100 microliter cryovials.

Non-COVID-19 plasma samples were collected from 60 adults originally recruited for a study of influenza vaccination response among the general population, at the

University of Georgia Clinical and Translational Research Unit (Athens, Georgia, USA) from September 2019 to December 2019. All volunteers were enrolled with written, informed consent. Participants were excluded if they, at the time of enrollment, already received the seasonal influenza vaccine, had acute or chronic conditions that would put the participant at risk for an adverse reaction to the blood draw or the flu vaccine (e.g., Guillain-Barré syndrome or allergies to egg products), or had conditions that could skew the analysis (e.g., recent flu symptoms or steroid injections/medications). Plasma samples analyzed in this present study were collected prior to vaccination.

Hospital-based autopsies for COVID-19 patients were performed at NYU Winthrop Hospital (Mineola, New York, USA) among persons with laboratory-confirmed COVID-19 or who were under investigation and tested positive on post-mortem PCR. Autopsies were performed between the dates of March 2020 and April 2020. The lungs, heart, kidneys, and liver were used in this study. Tissues were fixed in 10% buffered formalin for 72 h and routinely processed.

For information on ferret tissue collection please consult the complementary, publicly available study¹⁷⁹ by the Kelvin Lab.

Fluorescent Labelling of Samples

Total protein concentrations of plasma and tissue samples were determined with DC™ protein assay (Bio-Rad Laboratories). PBS refers to phosphate-buffered saline (137 mM NaCl, 2.7 mM KCl, 8.9 mM Na₂HPO₄, and 1.8 mM KH₂PO₄, pH = 7.4)

hereinafter. PBST refers to PBS supplemented with Tween® 20 (concentration in v/v indicated where it appears) hereinafter.

To label plasma proteins, each sample containing 10 µg total protein was first diluted in PBS to 27 µl. The pH of the solution was adjusted with 3 µL of 1M sodium bicarbonate. Then 0.21 µl of a 10 mg/ml Alexa Fluor™ 555 NHS ester (Thermo Fisher Scientific) stock solution was thoroughly mixed with the sample solution. The mixture was incubated in the dark and at room temperature with gentle agitation. After 1 hour, unconjugated dyes were removed by Zeba™ dye and biotin removal filter plates (Thermo Fisher Scientific). The reference standard, a commercial human plasma (Millipore Sigma, catalog #P9523), was fluorescently labelled with Alexa Fluor™ 647 NHS ester (Thermo Fisher Scientific) in a similar fashion. The amounts of reagents were scaled linearly to the starting protein amount (2 mg). Finally, each Alexa Fluor™ 555-labeled sample (10 µg of total protein) was mixed with a proper volume of Alexa Fluor™ 647-labeled reference standard containing the same amount of protein. The dual-color mixture was first diluted to 50 µl with PBS then mixed with 50 µl 0.1% PBST.

To label tissue samples from autopsy, each sample containing 50 µg total protein was first diluted in PBS to 60 µl. The pH of the solution was adjusted with 6.7 µL of 1M sodium bicarbonate. Then 0.2 µl of a 10 mg/ml Alexa Fluor™ 555 NHS ester (Thermo Fisher Scientific) stock solution was thoroughly mixed with the sample solution. The mixture was incubated in the dark and at room temperature with gentle agitation. After 1 hour, unconjugated dyes were removed by Zeba™ dye and biotin removal filter plates (Thermo Fisher Scientific). The pool reference was generated and fluorescently labelled

with Alexa Fluor™ 647 NHS ester (Thermo Fisher Scientific) in a similar fashion. Then, each Alexa Fluor™ 555-labeled sample (3 µg of protein) was mixed with a proper volume of Alexa Fluor™ 647-labeled reference standard containing the same amount of protein. The dual-color mixture was first diluted to 74 µl with PBS then mixed with 2 µl 0.2% PBST.

Prior to labelling, the formalin-fixed ferret lung tissues were incubated in 0.5M Tris (pH = 8.0) at 90 degrees Celsius for 1 hour. Then the tissues were thoroughly washed in PBS three times and homogenized in PBS using a bead ruptor. Homogenized tissue suspensions were incubated on ice for 1 hour, centrifuged at 18,000 x g for 30 minutes, and the resulting supernatants were kept for the next steps. To label the ferret tissue glycoproteins, each sample containing 10 µg total protein was first diluted in PBS to 27 µl. The pH of the solution was adjusted with 3 µL of 1M sodium bicarbonate. Then 0.2 µl of a 10 mg/ml Alexa Fluor™ 555 NHS ester (Thermo Fisher Scientific) stock solution was thoroughly mixed with the sample solution. The mixture was incubated in the dark and at room temperature with gentle agitation. After 1 hour, unconjugated dyes were removed by Zeba™ dye and biotin removal filter plates (Thermo Fisher Scientific). The pool reference was generated and fluorescently labelled with Alexa Fluor™ 647 NHS ester (Thermo Fisher Scientific) in a similar fashion. Then, each Alexa Fluor™ 555-labeled sample (3 µg of protein) was mixed with a proper volume of Alexa Fluor™ 647-labeled reference standard containing the same amount of protein. The dual-color mixture was first diluted to 74 µl with PBS then mixed with 2 µl 0.2% PBST.

Fabrication of Lectin Microarray Slides

Lectin microarray slides were fabricated as previously described¹²⁷ in a published protocol. In brief, lectins and antibodies were printed on Nexterion® Slide H (Applied Microarrays) with the microarray printer Nano-Plotter™ 2.1 (GeSim). The temperature and humidity inside the printer chamber were maintained at 14°C and 50%, respectively. Inhibiting sugars were added to lectin solutions to a final concentration of 50mM (except lactose: 25mM) prior to printing. Lectins for printing, concentrations and inhibiting sugars are listed in the appendices.

Dual-color Lectin Microarray

All steps were performed in the dark at room temperature. Each dual-color mixture was allowed to hybridize with the microarrays for 1 hour. Microarrays were washed twice with 0.005% PBST for 10 minutes and once with PBS for 5 minutes. The slides were briefly rinsed with ultrapure water and dried by centrifugation. Fluorescence signals were gained with Genepix™ 4400A fluorescence slide scanner (Molecular Devices) in the 532 nm channel and the 635 nm channel that correspond to the excitation/emission profiles of Alexa Fluor™ 555 and Alexa Fluor™ 647, respectively. Raw fluorescence signals and background signals were generated by the Genepix Pro™ 7 software (Molecular Devices), which were further processed and analyzed with a custom script as previously described⁸⁸. Heatmaps, boxplots and volcano plots were generated with R (Version: 4.0.1).

RNA-seq data analysis

Normalized RNA-seq data was retrieved from the complementary, publicly available study¹⁷⁹ by the Kelvin Lab. No further processing was needed prior to expression analysis.

Lectin Pulldown of Plasma Samples

In this section, centrifugation (1000x *g*, 2 minutes) was used to remove liquid from columns in washes and elutions. All steps were performed at room temperature. To prepare SNA-agarose columns, 200 µl 50% suspension of streptavidin-agarose resin (Millipore Sigma), was added to each microcentrifuge column. Storage buffer was removed, and the resin was washed with 200 µl PBS. 400 µl of biotinylated *Sambucus nigra* lectin (SNA, Vector Laboratories, pre-diluted to 0.5 mg/ml with PBS) was added to the column and the mixture was incubated with gentle agitation for 30 minutes. Then the resin was washed with 200 µl PBS twice. Control columns were prepared using the same procedure except that 400 µl of PBS was added to the column instead of biotinylated SNA.

To prepare SNA pulldown samples for mass spectrometry analysis, pooled plasma samples corresponding to the mild and severe COVID-19 patient group were prepared by combining equal volumes of individual samples. Each pooled plasma sample containing 300 µg of total protein was diluted to 300 µl with PBS. Pulldown was performed in triplicates (i.e., each pooled sample was enriched with three separate

columns prepared with the same procedure at the same time). Diluted samples were incubated with the SNA-bound resin or the control resin for 1 hour with gentle agitation. The resin was washed with 400 μ l PBS three times. To elute glycoproteins, 75 μ l of 0.2 M lactose in PBS was added to the column and incubated with gentle agitation. After 30 minutes, the flow-through was collected. Then 75 μ l of 0.2 M lactose in 0.2 M acetic acid was added to the column and incubated with gentle agitation. After 30 minutes, the flow-through was collected and combined with the previous flow-through. Finally, the pH of the combined eluate was adjusted with 1M Tris (pH = 9.0) to 7.5.

To prepare SNA pulldown samples for western blotting, pooled plasma samples corresponding to the mild COVID-19, severe COVID-19 and negative control group were prepared by combining equal volumes of individual samples. Albumin was depleted from each pooled sample with Pierce™ Albumin Depletion Kit (Thermo Fisher Scientific). Albumin-depleted sample protein concentrations were determined with *DC*™ protein assay (Bio-Rad Laboratories). Each albumin-depleted pooled plasma sample containing 200 μ g of total protein was diluted to 300 μ l with PBS. Diluted samples were incubated with the resin for 1 hour with gentle agitation. The resin was washed with 400 μ l PBS three times. To elute glycoproteins, 75 μ l of 0.2 M lactose in PBS was added to the column and incubated with gentle agitation. After 30 minutes, the flow-through was collected. Then 75 μ l of 0.2 M lactose in 0.2 M acetic acid was added to the column and incubated with gentle agitation. After 30 minutes, the flow-through was collected and combined with the previous flow-through. The combined eluate was then dialyzed against PBS.

Mass Spectrometry and Protein Identification

Protein quantitation was based on the number of peptide spectral matches (PSM). First, detected proteins (PSM ≥ 1 in at least one sample) were searched in the online portal of CRAPOME¹⁸³, a database of protein contaminants in proteomic experiments. CRAPOME outputs a ratio of [Num of Expt. (found/total)] for each query protein. Any protein with a [Num of Expt. (found/total)] > 0.2 is considered a contaminant and removed.

To identify non-specifically binding proteins, two tailed student's t-test was performed between the PSM of the proteins in the triplicates of the pulldown samples and corresponding bead-only controls (PSM_{PD} and PSM_{CT}, respectively). Any protein that satisfies 1) average PSM_{PD} \leq average PSM_{CT}, or 2) PSM_{PD} < 2 , or 3) p-value of the t-test > 0.05 , is removed.

To identify significantly upregulated proteins in SNA-enriched severe COVID-19 plasma, two tailed student's t-test was performed between the PSM of the enriched proteins in the triplicates of the severe sample and the mild sample (PSM_{severe} and PSM_{mild}, respectively). Any protein that satisfies 1) average PSM_{severe} $>$ average PSM_{mild} and 2) p-value of the t-test < 0.05 is considered significantly upregulated in severe COVID-19 plasma.

Western Blotting

All steps were conducted at room temperature unless noted otherwise. The column eluate, or the corresponding input (albumin-depleted plasma) containing 20 µg of total protein, was mixed with Laemmli buffer to a final volume of 200 µl. Then 100 µl of each sample was heated at 90°C before resolved by 4-20% SDS-PAGE. Proteins were transferred to a nitrocellulose membrane, which was then stained with Ponceau S. After the total protein stain was erased, the membrane was blocked with a blocking buffer (PBS with 3% (w/v) BSA and 0.05% (v/v) Tween® 20, pH = 7.4) for 1 hour. Then the membrane was incubated with primary antibodies pre-diluted to 1 µg/ml in the blocking buffer for 1 hour. Rabbit anti-human complement C5 antibody (clone# EPR19699-24, Abcam, catalog# ab202039) and mouse anti-human complement C9 antibody (clone# X197, Hycult Biotech, catalog# HM2111) were used for C5 and C9 detection, respectively. The membrane was washed with 0.05% PBST three times for 5 minutes per wash then incubated with secondary antibodies pre-diluted to 0.1 µg/ml in the blocking buffer for 15 minutes. CF™640-conjugated, goat anti-rabbit IgG antibody (Millipore Sigma, catalog# SAB4600399) and IRDye® 800CW-conjugated, goat anti-mouse IgG antibody (LI-COR, catalog# 926-32210) were used for C5 and C9 primary antibody detection, respectively. Finally, the membrane was washed with 0.05% PBST three times for 5 minutes per wash before imaging.

Chapter 4. Discovery of Glycan Biomarkers Predicting Vaccine Responses

4.1 Introduction

Vaccines remain the most effective tool to control the spread and the severity of infectious diseases in the human population.¹⁸⁴ For example, it was estimated that vaccination against influenza, a common respiratory infectious disease killing 444,000-553,000 people annually worldwide, reduces the risk of infections that need medical attention by 50%.^{185–187} According to the estimate of the US Centers for Disease Control and Prevention, influenza vaccination prevented over 105,000 hospitalizations and 7.5 million influenza illnesses in the US during the 2019–2020 influenza season.¹⁸⁸

The responses to influenza vaccines, however, can be highly variable depending on host factors.¹²⁶ Well-established host factors that impact influenza vaccine effectiveness are primarily descriptive characteristics, such as sex, age, and body-mass index (BMI). Advances in the past few decades have found a number of host factors affecting vaccine responses at molecular level, including hormones, pro-inflammatory cytokines, pre-existing antibodies, pre-existing T cell compositions, and polymorphisms of certain genes.^{189–192} Identification of these biomarkers is instructive for improving vaccines and vaccination strategy. For instance, low-grade inflammation, characterized by increased baseline levels of pro-inflammatory mediators, is often implicated in ageing and obesity.^{193,194} This inflammation is believed to suppress immune function and thus reduce vaccine effectiveness.¹⁹⁵ Based on this, anti-inflammatory drugs have been used in some studies to improve vaccine response. One study showed metformin, a diabetes

drug with anti-inflammatory potency, increased antibody responses to influenza vaccines in obese, diabetic patients through suppressing B cell-mediated inflammation.¹⁹⁶

In the search for clinical biomarkers associated with vaccine responses, however, glycans have been overlooked, despite the critical roles of glycosylation in maintaining the functionality of the immune system (see Chapter 1). Prior to the present study, only three studies have looked into the associations between host glycosylation and vaccine response in human subjects, the major findings of which are summarized in

Table 4.1.

Table 4.1 Associations of host glycosylation and vaccine response in influenza-vaccinated human cohorts from past and present studies.

Study	Cohort	Location	Main Finding(s)
Selman et al. (2012) ¹⁹⁷	Healthy adults (n = 10)	The Netherlands	Increased galactosylation and sialylation of influenza-specific plasma IgG post vaccination
	Healthy children (n = 10)	Gabon	Increased galactosylation and sialylation of influenza-specific plasma IgG post vaccination
Wang, J.R. et al. (2015) ¹⁹⁸	Healthy adults (n = 26)	China	High mannose glycans and some truncated (monoantennary) glycans of total serum IgG1 are higher in responders to influenza B
Wang, T.T. et al. (2015) ¹⁹⁹	Healthy adults (n = 10)	USA	Higher sialylation of anti-hemagglutinin IgG Fc chain is associated with higher antigen affinity
This study	Healthy adults (n = 160)	USA	See below

As shown in **Table 4.1**, previous studies have some major limitations. First, they have primarily focused on the *N*-glycosylation of IgG, likely due to the better characterized roles of glycosylation in IgG (see Chapter 1) and relatively simple glycosylation profile (conserved, single *N*-glycosylation site in Fc domain; poorly *O*-glycosylated).^{200,201} Second, association of vaccine response with pre-vaccination glycosylation was not examined. Thirdly, the cohort sizes were small (all < 30 subjects).

In the present study described in this chapter, I analyzed the pre- and post-vaccination glycomes of the serum samples from 160 adults receiving a commercial quadrivalent influenza vaccine and associated the glycosylation profiles with the differential antibody responses in this cohort. Hitherto, this is the largest-scale clinical study on the association between vaccination and glycosylation. Moreover, this study is focuses on both the *N*- and *O*-glycosylation of whole serum, rather than just IgG, thus broadening the scope of previous studies. In brief, I identified the blood group epitope Lewis A (Le^a) as a pre-vaccination biomarker for unresponsiveness to influenza vaccine. Glycoproteomic analysis showed many serum glycoproteins can carry Le^a, especially proteins in the complement pathway. Post-vaccination changes in select glycans (sialyl Lewis X, sLe^x; high mannose) were found only in individuals who responded strongly to the vaccine, indicative of active immune remodeling post-vaccination. By unifying the findings from this study and previous literature, I also propose potential mechanisms of how glycan biomarkers biologically associate with immune response.

4.2 Results and Discussion

4.2.1 Pre-vaccination serum Le^a differentiates high and non-responders

(Note: in this section, antibody titer determination was performed by collaborators including Michael Carlock)

The cohort was comprised of 160 Caucasian, non-smoker adult volunteers who were given the Fluzone™ influenza vaccine during the 2019-2020 flu season at the University of Georgia Clinical and Translational Research Unit (**Table 4.2**). The quadrivalent vaccine used in this study is composed of inactivated split virions corresponding to two influenza virus A strains and two influenza virus B strains: A/Brisbane/02/2018 (subtype H1N1), A/Kansas/14/2017 (subtype H3N2), B/Phuket/3073/2013 (Yamagata lineage), and B/Colorado/6/2017-like (Victoria lineage).

Volunteer serum samples were collected on the day of vaccination (*pre-vaccination; d0*) and approximately four weeks after vaccination (*post-vaccination; d28*). Antibody response in this study was determined based on pre- and post-vaccination anti-hemagglutinin (anti-HA) antibody titers measured with hemagglutination inhibition (HAI) assays for each strain. Response score is defined as the adjusted logarithm of the fold change in titers (d28 vs. d0), which serves as a metric of antibody response to each strain (see Methods for details). A total response score was then calculated by summing the strain-specific response scores. Based on the overall scores, volunteers were categorized into high responders (total response score ≥ 8), low/moderate responders

($4 \leq$ total response score < 8), and non-responders (total response score < 4) (**Table 4.2**).

Table 4.2. Characteristics of Study Participants (n=160). Reprinted with permission from *J. Proteome Res.* DOI: 10.1021/acs.jproteome.2c00251. Copyright 2022 American Chemical Society.

	High responders (n = 66)	Low/moderate responders (n = 39)	Non responders (n = 54)
Male/Female	22/44	15/24	24/30
Median Age in Years (IQR)	46.5 (31.0-58.8)	56 (39.9-67.0)	56.5 (42.3-70.0)
Median Body-Mass Index (IQR)	28.7 (25.2-32.8)	30.8 (27.4-34.1)	26.9 (24.4-31.0)

Lectin microarray-based glycomic profiling was performed to investigate whether the pre-vaccination serum glycomes vary with responder categories.⁸⁷ A list of probes printed on the microarray slides can be found in the appendices. Heatmap presentation of the microarray data of the pre-vaccination sera is shown in **Figure 4.1**.

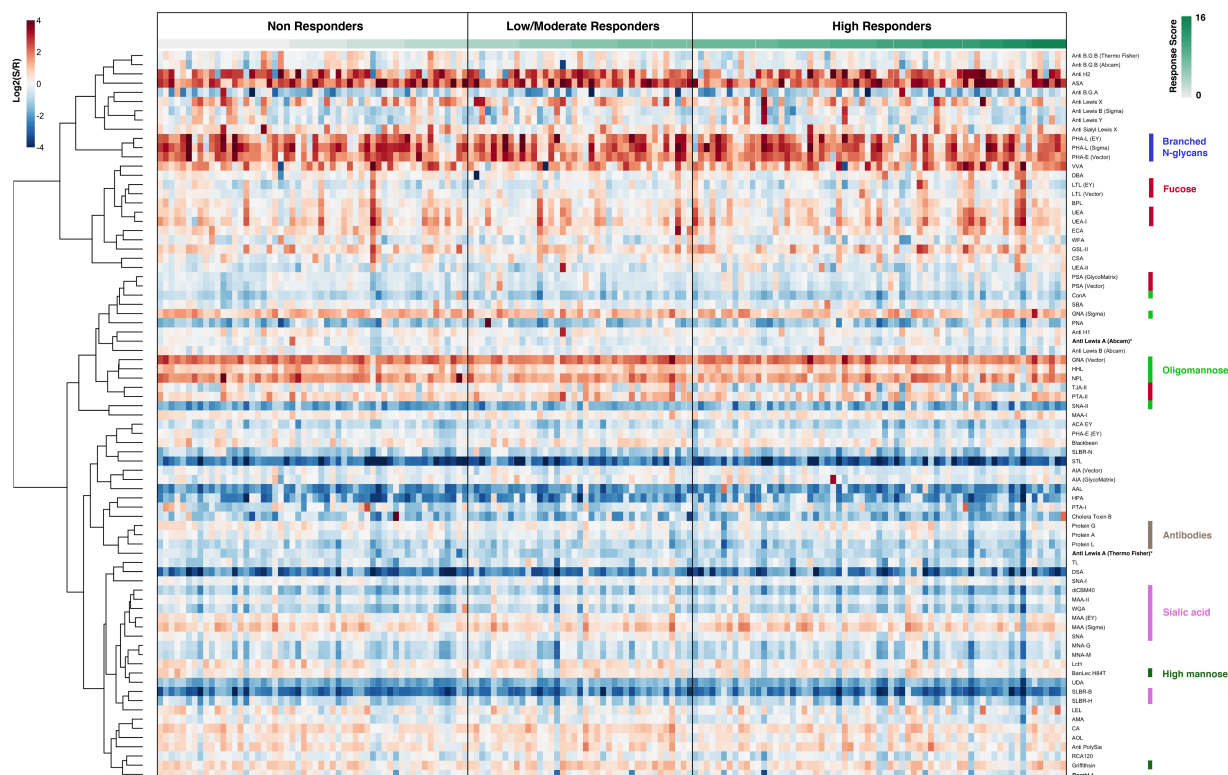


Figure 4.1. Heatmap of lectin microarray data for d0 serum samples. Columns represent the participants and rows represent the probes. Color of cells represent the normalized \log_2 ratios (Sample signal (S)/Reference signal (R)). Total response scores are annotated with a green-white sliding scale bar on the top of the heatmap. Reprinted with permission from *J. Proteome Res.* DOI: 10.1021/acs.jproteome.2c00251. Copyright 2022 American Chemical Society.

I first compared the pre-vaccination serum glycosylation data of high responders and non-responders, as the two groups made up most (~76%) of the cohort and they were cleanly separated by antibody responses. As shown in **Figure 4.2a**, non-responder sera had higher levels of fucosylated, type I LacNAc antigens with respect to high responders (BambL, anti-Le^a, anti-H1). Most significantly, the blood group antigen Lewis A (Le^a) was elevated in non-responders, with three independent probes showing the same trend (BambL, two anti-Le^a).²⁰² Type I blood group antigen H (H1 antigen) was also higher in non-responders (anti-H1). H1 antigen is biosynthetically associated

with Le^a (Figure 4.2b), possibly indicating an overall increase in the type I LacNAc motif.

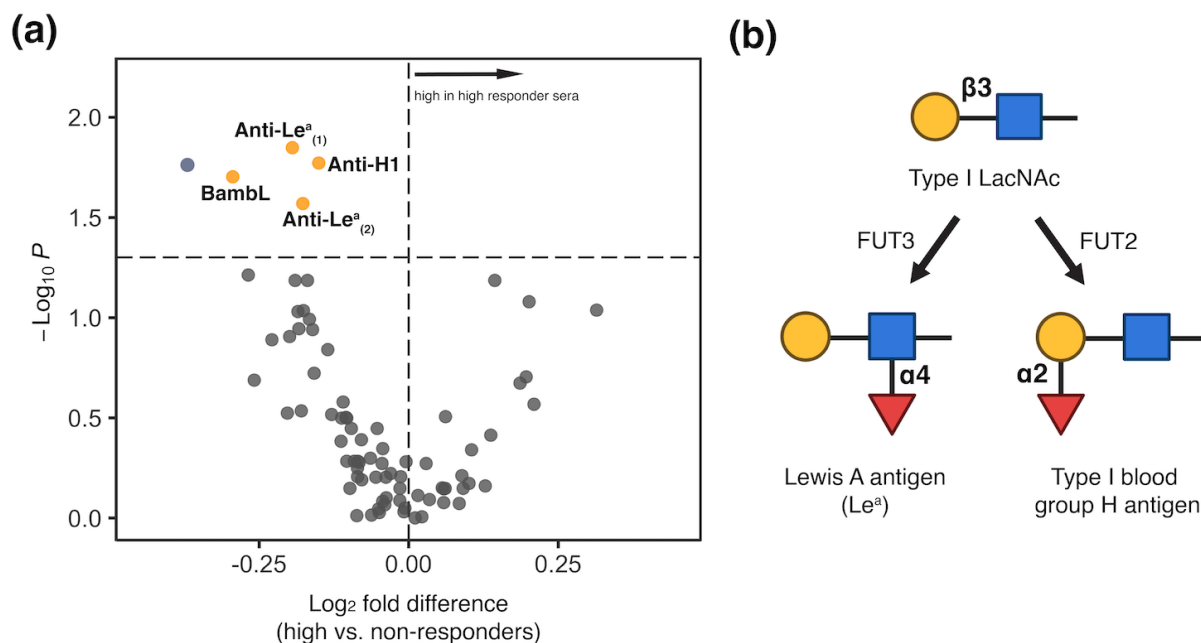


Figure 4.2. (a) Volcano plot comparing lectin microarray data for high responders ($N = 65$) and non-responders ($N = 54$) pre-vaccination. Mann–Whitney U test was used to determine p -values. Probes with $p < 0.05$ are colored in yellow. (b) Partial biosynthetic routes of Lewis A antigen and type I blood group antigen H (O) antigen. FUT2: galactoside alpha-(1,2)-fucosyltransferase 2; FUT3: 3-galactosyl-N-acetylglucosaminide 4-alpha-L-fucosyltransferase. Reprinted with permission from *J. Proteome Res.* DOI: 10.1021/acs.jproteome.2c00251. Copyright 2022 American Chemical Society.

Antibody response to influenza vaccines is influenced by multiple factors such as age and BMI. In my collaborators' recent work among the same cohort, machine learning tools were utilized to construct a multiple linear regression (MLR) model that estimates the effects of multiple factors (sex, age, BMI, pre-vaccination antibody titers, etc.) on antibody response, and adjusted antibody response scores (MLR-adjusted scores) were generated after correcting for those effects.²⁰³ Using the MLR-adjusted

scores, I reclassified the cohort (high responders: top tercile scores; non-responders: bottom tercile scores) and performed the same comparison between high responders and non-responders. Higher serum Le^a in non-responders was again observed (**Figure 4.3**), suggesting this glycan motif is associated with the “clean” antibody response that is not influenced by common confounding factors.

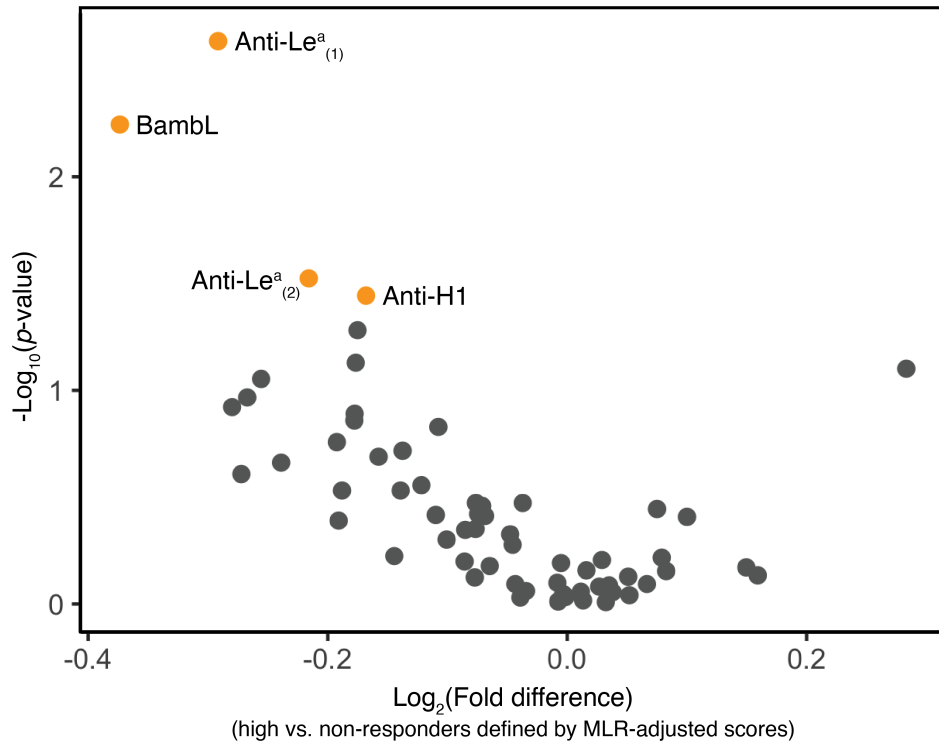


Figure 4.3. Volcano plot comparing pre-vaccination lectin microarray data for high responders (n = 48) and non-responders (n = 48), re-classified using the MLR-adjusted scores with estimated impacts of multiple confounding factors removed (see text and reference²⁰³). A negative fold change denotes higher binding of the probe in non-responders. Mann–Whitney U test was used to determine p-values. Probes with $p < 0.05$ are labelled and colored in yellow.

Next, I compared the pre-vaccination serum glycosylation data of high-responders and low/moderate responders. Many individuals in the latter subgroup

responded strongly to one single strain of influenza but poorly to other strains. Despite this, one Le^a-binding probe (BambL) still exhibited lower binding to low/moderate responder sera compared to high responders (**Figure 4.4**). This suggested that evidence of association between higher Le^a and low vaccine response can still be observed even between responder groups that were not distinctly separated.

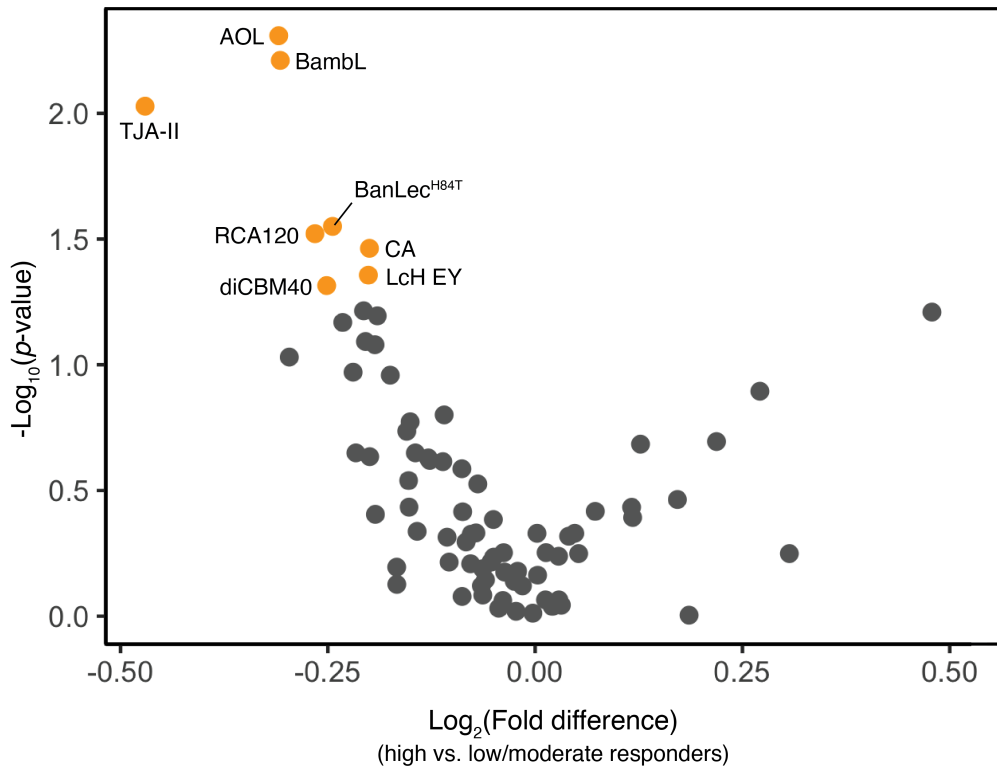


Figure 4.4. Volcano plot comparing lectin microarray data for high responders (n = 66) and low/moderate-responders (n = 39) pre-vaccination. A negative fold change denotes higher binding of the probe in non-responders. Mann–Whitney U test was used to determine p-values. Probes with $p < 0.05$ are colored in yellow.

Responses to influenza vaccination and infection vary with influenza strains.^{204–}

²⁰⁶ Therefore, one can assume that host factors associate with influenza vaccine response differently, depending on the strains. To investigate the association of pre-

vaccination serum glycosylation and antibody response in a strain-specific manner, I compared the corresponding microarray data between the by-strain high-responders and non-responders (high responder: strain score ≥ 2 , non-responder: strain score < 1). For three out of the four strains, higher serum Le^a was observed in non-responders compared to high responders (Conventional, **Table 4.3**), in concordance with the results of the comparison using total scores. Similar results were obtained when comparing the high-responders and low/moderate responders reclassified based on the MLR-adjusted scores (high responders: top tercile scores; low/moderate responders: middle tercile scores) (MLR, **Table 4.3**). Summarizing, I discovered pre-vaccination serum Le^a as a glycan biomarker of antibody response to influenza vaccination, with people lacking responses exhibiting lower levels of pre-vaccination Le^a. The association does not seem to change with factors such as sex, age, and BMI, and does not vary significantly with the influenza strains.

Table 4.3. Differences in pre-vaccination microarray data of Le^a-binding lectins between high responders and non-responders to each influenza strain.

Strain	Method used to define responders	Anti-Le ^a ₍₁₎	Anti-Le ^a ₍₂₎	BamBL
		Log ₂ [Fold difference between high and non-responders] (<i>p</i> -value)		
A/Brisbane/02/2018 (H1N1)	Conventional	-0.18 (0.021)	-0.17 (0.015)	-0.30 (0.004)
	MLR	-0.29 (0.009)	-0.09 (0.244)	-0.48 (0.001)
A/Kansas/14/2017 (H3N2)	Conventional	-0.04 (0.222)	-0.12 (0.094)	-0.20 (0.048)
	MLR	-0.08 (0.182)	-0.09 (0.391)	-0.16 (0.155)

B/Phuket/3073/2013 (Yamagata)	Conventional	-0.09 (0.361)	-0.21 (0.028)	-0.16 (0.221)
	MLR	-0.29 (0.015)	-0.24 (0.035)	-0.31 (0.040)
B/Colorado/6/2017- like (Victoria)	Conventional	-0.11 (0.196)	-0.08 (0.220)	-0.18 (0.118)
	MLR	-0.31 (0.005)	-0.08 (0.348)	-0.27 (0.035)

Note: (1) p -values less than 0.05 are bolded; (2) A negative fold change denotes higher binding of the probe in non-responders; (3) See Methods (section 4.4) for details of defining responder categories using conventional response scores and MLR-adjusted response scores.

4.2.2 Serum Le^a is enriched in complement proteins

As glycosylation may differ with glycoproteins in serum, I next performed Le^a-focused glycoproteomic analysis. This may provide insights into the potential mechanisms linking high serum Le^a to low antibody response. Anti-Le^a and BambL were used to pull down a pooled serum from all volunteer samples, and compositions of the enriched glycoproteins were resolved by mass spectrometry (**Figure 4.5a**). In total, 79 glycoproteins were pulled down by BambL, and 30 were pulled down by anti-Le^a (**Figure 4.5b**; full list in **Table 4.4**). Consistent with the broader specificity profile, BambL enriched a wider range of glycoproteins compared to anti-Le^a. **Table 4.4** showed that a variety of serum proteins can bear Le^a, including antibodies, proteins of the complement pathways, cell adhesion molecules, protease inhibitors, and blood coagulation factors. To see the functional commonality of glycoproteins carrying Le^a, I performed Gene Ontology pathway enrichment analysis. As shown in **Figure 4.5c**, complement activation pathways and humoral immune pathways were among the top enriched pathways. Moreover, most proteins hits were found in the complement

activation pathways, indicating strong representation of Le^a in complement pathway glycoproteins.

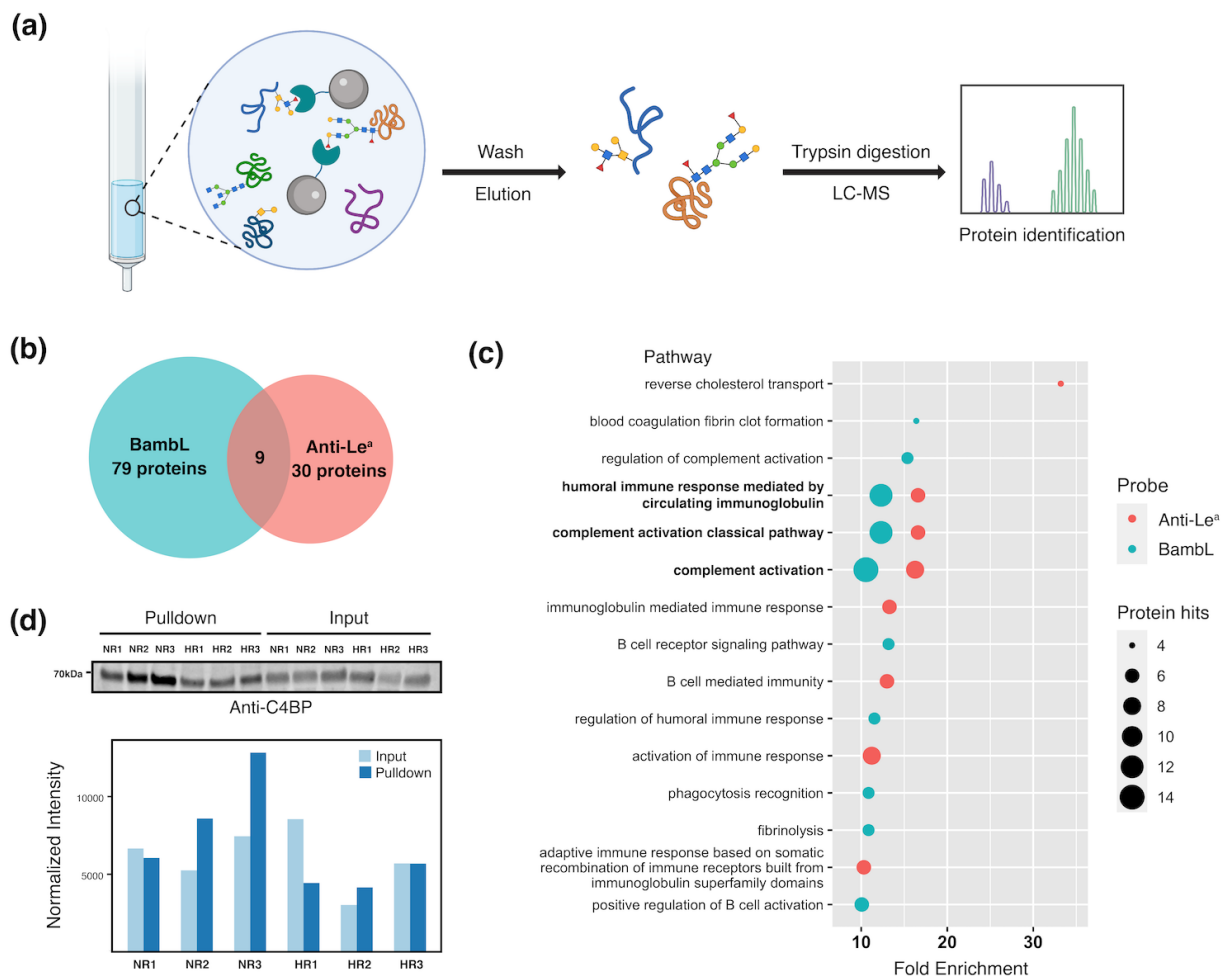


Figure 4.5. (a) Scheme of the experimental approach of glycoproteomic analysis. (b) Number of glycoproteins identified in BamBL/anti-Le^a pull-down experiments. (c) Gene ontology pathway enrichment analysis for glycoproteins enriched with BamBL/anti-Le^a. The false discovery rates (FDRs) of the enriched pathways shown are all < 0.05. (d) Differential C4BP glycosylation. Western blot analysis for C4BP of BamBL pull-down samples and corresponding input for three high responders (HR1, HR2 and HR3) and three non-responders (NR1, NR2, and NR3) is shown. Signal intensities of the bands (normalized to total protein stain) are depicted in the bar plot. Reprinted with permission from *J. Proteome Res.* DOI: 10.1021/acs.jproteome.2c00251. Copyright 2022 American Chemical Society.

Table 4.4. Serum glycoproteins enriched by BambL and/or anti-Le^a

Uniprot Accession	Protein name	BambL	Anti-Le^a
Q9HDC9	Adipocyte plasma membrane-associated protein	+	
P43652	Afamin		+
P08697	Alpha-2-antiplasmin		+
P02765	Alpha-2-HS-glycoprotein	+	+
P01023	Alpha-2-macroglobulin	+	
P12821	Angiotensin-converting enzyme	+	
P02656	Apolipoprotein C-III		+
P05090	Apolipoprotein D	+	
O95445	Apolipoprotein M	+	+
P08519	Apolipoprotein(a)	+	+
O75882	Attractin	+	
P02749	Beta-2-glycoprotein 1	+	+
P04003	C4b-binding protein alpha chain	+	+
P12830	Cadherin-1	+	
P55290	Cadherin-13	+	
P33151	Cadherin-5	+	
Q96IY4	Carboxypeptidase B2	+	
Q9NQ79	Cartilage acidic protein 1		+
P35222	Catenin beta-1		+
P07858	Cathepsin B		+
Q6YHK3	CD109 antigen	+	
Q15762	CD226 antigen	+	
P16070	CD44 antigen	+	
Q9BY67	Cell adhesion molecule 1	+	
P43121	Cell surface glycoprotein MUC18	+	
P11597	Cholesteryl ester transfer protein		+
P06276	Cholinesterase		+
P10909	Clusterin	+	+
P12259	Coagulation factor V	+	+
Q9NZP8	Complement C1r subcomponent-like protein		+

P09871	Complement C1s subcomponent	+	
P01024	Complement C3	+	
P07357	Complement component C8 alpha chain		+
P02748	Complement component C9		+
P00751	Complement factor B		+
P08603	Complement factor H	+	
P05156	Complement factor I	+	
Q02413	Desmoglein-1		+
P17813	Endoglin	+	
Q9UNN8	Endothelial protein C receptor	+	
P02671	Fibrinogen alpha chain	+	
P02675	Fibrinogen beta chain	+	
P02751	Fibronectin		+
P23142	Fibulin-1	+	
Q15485	Ficolin-2		+
Q92820	Gamma-glutamyl hydrolase		+
P05546	Heparin cofactor 2	+	
P26927	Hepatocyte growth factor-like protein	+	
Q14520	Hyaluronan-binding protein 2	+	
P01876	Immunoglobulin heavy constant alpha 1	+	
P01877	Immunoglobulin heavy constant alpha 2	+	
P01857	Immunoglobulin heavy constant gamma 1	+	
P01859	Immunoglobulin heavy constant gamma 2	+	
P0DP01	Immunoglobulin heavy variable 1-8		+
P15814	Immunoglobulin lambda-like polypeptide 1	+	
P17936	Insulin-like growth factor-binding protein 3	+	
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	+	+
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	+	
P13598	Intercellular adhesion molecule 2	+	
P14151	L-selectin	+	
Q8N6C8	Leukocyte immunoglobulin-like receptor subfamily A member 3	+	
P08637	Low affinity immunoglobulin gamma Fc region receptor III-A	+	
P51884	Lumican	+	+
P13473	Lysosome-associated membrane glycoprotein 2	+	

P07333	Macrophage colony-stimulating factor 1 receptor	+
P48740	Mannan-binding lectin serine protease 1	+
O00187	Mannan-binding lectin serine protease 2	+
P10721	Mast/stem cell growth factor receptor Kit	+
P01033	Metalloproteinase inhibitor 1	+
Q7Z5P9	Mucin-19	+
P05164	Myeloperoxidase	+
Q96PD5	N-acetylmuramoyl-L-alanine amidase	+
P13591	Neural cell adhesion molecule 1	+
O00533	Neural cell adhesion molecule L1-like protein	+
Q99784	Noelin	+
P10153	Non-secretory ribonuclease	+
P16109	P-selectin	+
Q6UXB8	Peptidase inhibitor 16	+
P62937	Peptidyl-prolyl cis-trans isomerase A	+
P05154	Plasma serine protease inhibitor	+
P00747	Plasminogen	+
P07359	Platelet glycoprotein Ib alpha chain	+
P40197	Platelet glycoprotein V	+
Q6UX71	Plexin domain-containing protein 2	+
P15151	Poliovirus receptor	+
P01833	Polymeric immunoglobulin receptor	+
P20742	Pregnancy zone protein	+
Q8NBP7	Proprotein convertase subtilisin/kexin type 9	+
P41222	Prostaglandin-H2 D-isomerase	+
Q9UK55	Protein Z-dependent protease inhibitor	+
Q12913	Receptor-type tyrosine-protein phosphatase eta	+
P10586	Receptor-type tyrosine-protein phosphatase F	+
P35542	Serum amyloid A-4 protein	+
P02743	Serum amyloid P-component	+
Q15166	Serum paraoxonase/lactonase 3	+
P22105	Tenascin-X	+
P05452	Tetranectin	+
P02786	Transferrin receptor protein 1	+

P35916	Vascular endothelial growth factor receptor 3	+
P02774	Vitamin D-binding protein	+

Note: “+” indicates that the protein was detected in the sample enriched by the corresponding probe.

Among the glycoproteins enriched by both BamBL and anti-Le^a, C4BP is a relatively high abundance serum glycoprotein (~ 0.2 mg/ml),²⁰⁷ thus it might have contributed significantly to the observed differences in lectin binding. C4b-binding protein (C4BP) was then selected to confirm that glycosylation of serum proteins varies between samples. For this validation, three non-responder samples (NR1, NR2, NR3) showing high binding to BamBL and three high responder samples (HR1, HR2, HR3) showing low binding to BamBL were separately enriched with BamBL. Anti-C4BP Western blotting was then performed on the input samples and the corresponding enriched samples (**Figure 4.5d**). In all six samples, BamBL enriched C4BP. Although the total C4BP levels varied among the samples, enrichment of C4BP by BamBL was obvious, and the degree of enrichment (pulldown/input ratio) also varied, confirming differential C4BP glycosylation among the samples.

4.2.3 Potential biological links between Le^a and vaccine response

Genetics is a prominent determinant of Le^a expression in populations. In the presence of functional galactoside alpha-(1,2)-fucosyltransferase 2 (FUT2), conversion from type I LacNAc to type I H antigen takes place much faster than to Le^a (**Figure 4.2b**).⁹⁹ Therefore, individuals with functional deficiency in FUT2 (known as “non-secretors”) tend to secrete higher levels of Le^a-bearing glycoproteins into bodily fluids,

compared to people with functional FUT2 (known as “secretors”). Non-secretors represent approximately 20% of the population.^{99,208,209} Therefore, any impact of Le^a may have on vaccine effectiveness will be on a large scale.

The sole characterized role of Le^a in immunity is influencing the susceptibility to infections and disease severity.^{99,210,211} For pathogens such as rotavirus and norovirus that express blood group antigen-specific lectins for attachment to host cells, differential glycan binding can in part explain their infectivity varying with host Le^a level/secretor status. Resembling the finding in the present study, lower levels of serum Le^a were associated with weaker antibody responses in children vaccinated against rotavirus, which may be attributed to differential glycan binding.²¹² This mechanism involving glycan binding does not apply to pathogens devoid of affinity to blood group epitopes such as influenza virus. Non-secretors (high Le^a) are less susceptible to influenza infection,²¹³ hinting at a formerly uncharacterized role of Le^a in immune response to influenza antigens.

An emerging mechanism of how Le^a impacts vaccine response is through shaping the microbiota (**Figure 4.6**).^{99,214} Microbiome can secrete immunomodulatory metabolites, sense antigens, and reprogram immune cells, all of which influence host immunity.²¹⁵ Multiple studies have shown secretors and non-secretors have distinct microbiome compositions. Non-secretor microbiota are more abundant in certain bacteria including the genus *Bacteroides*,^{211,216,217} which is negatively associated with antibody response to influenza vaccines in pigs.²¹⁸ Findings reported by Xiao et al. further strengthen the link between Le^a and vaccine response. They found response to

influenza vaccine in mice was ameliorated by feeding 2'-fucosyllactose, which mimics the α 1-2 fucose motif in type I H antigen and Lewis B antigen that are lower in non-secretors.²¹⁹ In aggregate, multiple studies have pointed to a role of the host microbiota in determining the vaccine response in secretors and non-secretors, although the exact mechanisms are still elusive.

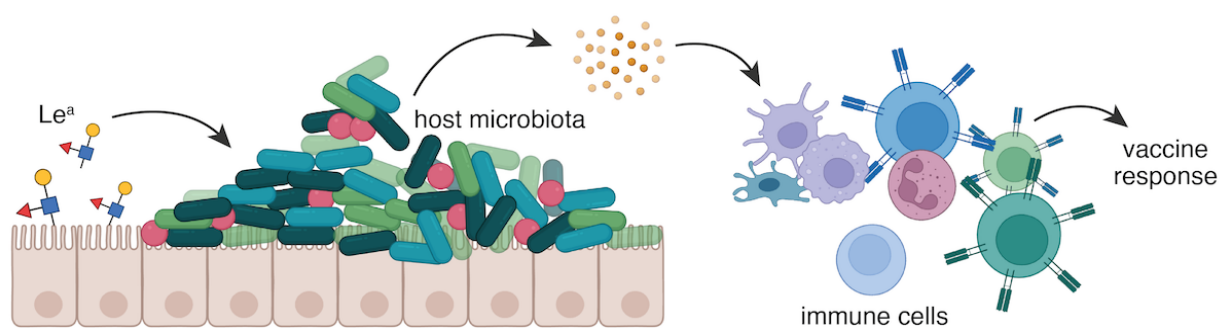


Figure 4.6. A possible mechanism of how Le^a levels impacts vaccine response. Relative abundance of Le^a or other secretor/non-secretor glycans in the host environment shapes the host microbiota, which regulates the functions and activity of immune cells (probably via release of immunomodulatory molecules) and thus influences response to vaccines.

In this study, Le^a was enriched in a number of proteins in the complement activation pathway. Several studies have shown the complement system is important in eliciting vaccine responses. C3 knockout mice exhibit weak antibody response to both influenza infection and vaccination, and human C2 deficiency is correlated with low antibody response to pneumococcal vaccines.^{220,221} One study has found greater lectin-mediated complement activation triggered by salivary glycoproteins in secretors compared to non-secretors.²²² As discussed in Chapter 1 and Chapter 3, the functional significance of glycosylation of complement proteins themselves is poorly understood.

The present study has suggested a potential role of complement protein glycosylation in mediating antibody responses to vaccines.

4.2.4 Post-vaccination glycomic changes in high responders indicate active immune response

Whether changes in glycosylation can be observed at whole serum level as a response to vaccination was unknown. To investigate this, I profiled the post-vaccination serum glycomes with lectin microarrays (**Figure 4.7a**), and compared their corresponding pre-vaccination glycomes. Overall, pre- and post-vaccination glycomic data from the same volunteers were highly correlated, regardless of responder categories (**Figure 4.8**). This indicates the overall serum glycomes were relatively stable in response to vaccination, in line with findings from previous studies.^{197,198}

When I examined binding data of individual lectins, some differences in glycosylation changes were observed between high responders and non-responders. In high responders, high mannose glycans were downregulated after vaccination, which was not observed in non-responders (**Figure 4.7b**). High mannose glycans bind mannose binding lectin (MBL), an upstream trigger of the lectin-mediated complement activation. Recent studies have shown MBL and complement are needed to increase response to influenza antigens in mice.^{223,224} High mannose may be biologically associated with vaccine response via this mechanism.

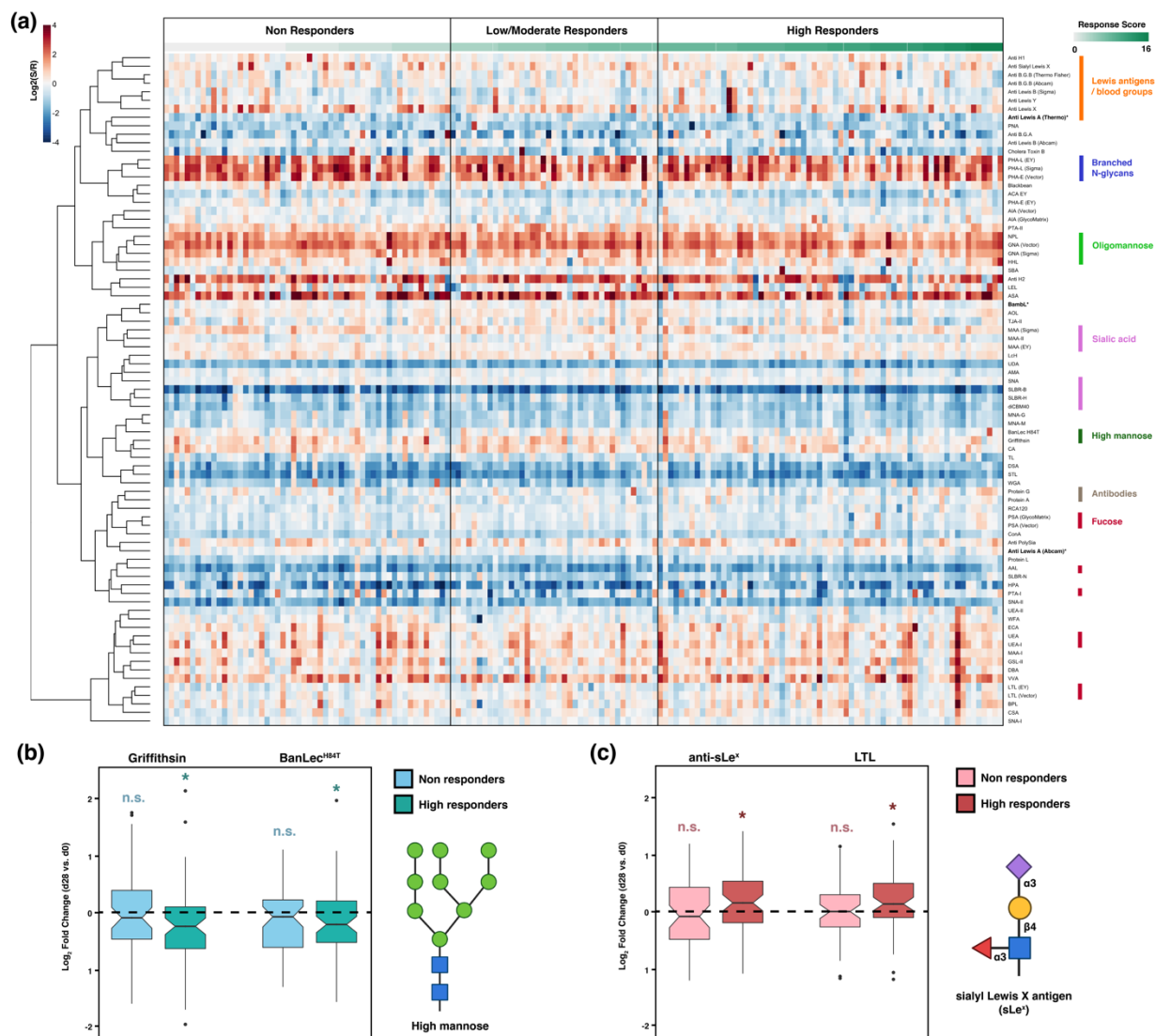


Figure 4.7. (a) Heatmap of lectin microarray data for post-vaccination (day 28) serum samples. Columns represent the participants and rows represent the probes. Shown is the normalized \log_2 ratios (Sample signal (S)/Reference signal (R)). Total response scores are annotated with a green-white sliding scale bar on the top of the heatmap. Rough specificities of select lectins are annotated on the right of the heatmap. (b, c) Boxplots of \log_2 Fold-Change for paired d28 and d0 samples in non- and high-responders. (b) High-mannose binding lectins (Griffithsin and BanLec^{H84T}). (c) sialyl Lewis X binding probes (Anti-sLe^x). Paired Mann–Whitney U test was used to determine p -values. n.s.: not statistically significant (no difference is observed d28/d0); (*) $p < 0.05$. Reprinted with permission from *J. Proteome Res.* DOI: 10.1021/acs.jproteome.2c00251. Copyright 2022 American Chemical Society.

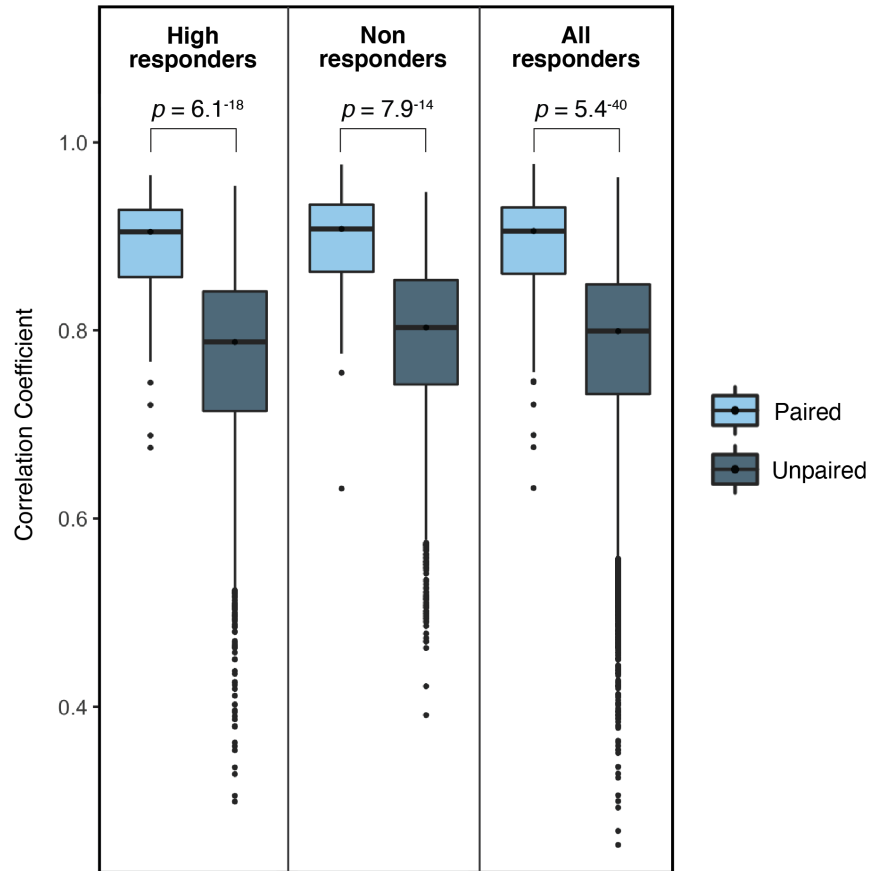


Figure 4.8. Boxplots comparing correlation coefficients between pre- and post-vaccination lectin microarray data. Paired: correlation coefficients of the pre- and post-vaccination glycomes of the same individuals; Unpaired: correlation coefficients of the pre- and post-vaccination glycomes of different individuals. Student's t-test was performed on the Fisher transformed-correlation coefficients.

In high responders, I also observed upregulation of multiple glycan motifs bearing fucose, including Lewis B antigen (anti-Le^b), Lewis X antigen (anti-Le^x), sialyl Lewis X antigen (anti-sLe^x), α1,3-fucose (LTL), and pan-fucose (AAL) (**Figure 4.7c**; **Figure 4.9a**). This suggests a general upregulation in total fucosylation in response to vaccination. It has been observed in multiple systems that infection induces fucosylation and expression of fucosyltransferases.^{76,225–229} Thus, increased binding of post-vaccination sera exclusively in high responders indicated active immune response in

this subgroup. In contrast, non-responders did not exhibit significant changes in serum glycosylation after immunization, but there appeared to be a slight decrease in total antibodies (protein A, protein G, protein L, **Figure 4.9b**). When directly comparing the post-vaccination glycomes of high and non-responders, no difference in Le^a was observed (**Figure 4.10**). This is most likely because vaccination augmented total fucosylation, which altered the Le^a pool. Longitudinal studies over a longer time span are needed to find out whether and when the serum glycans that were altered after vaccination return to pre-vaccination levels.

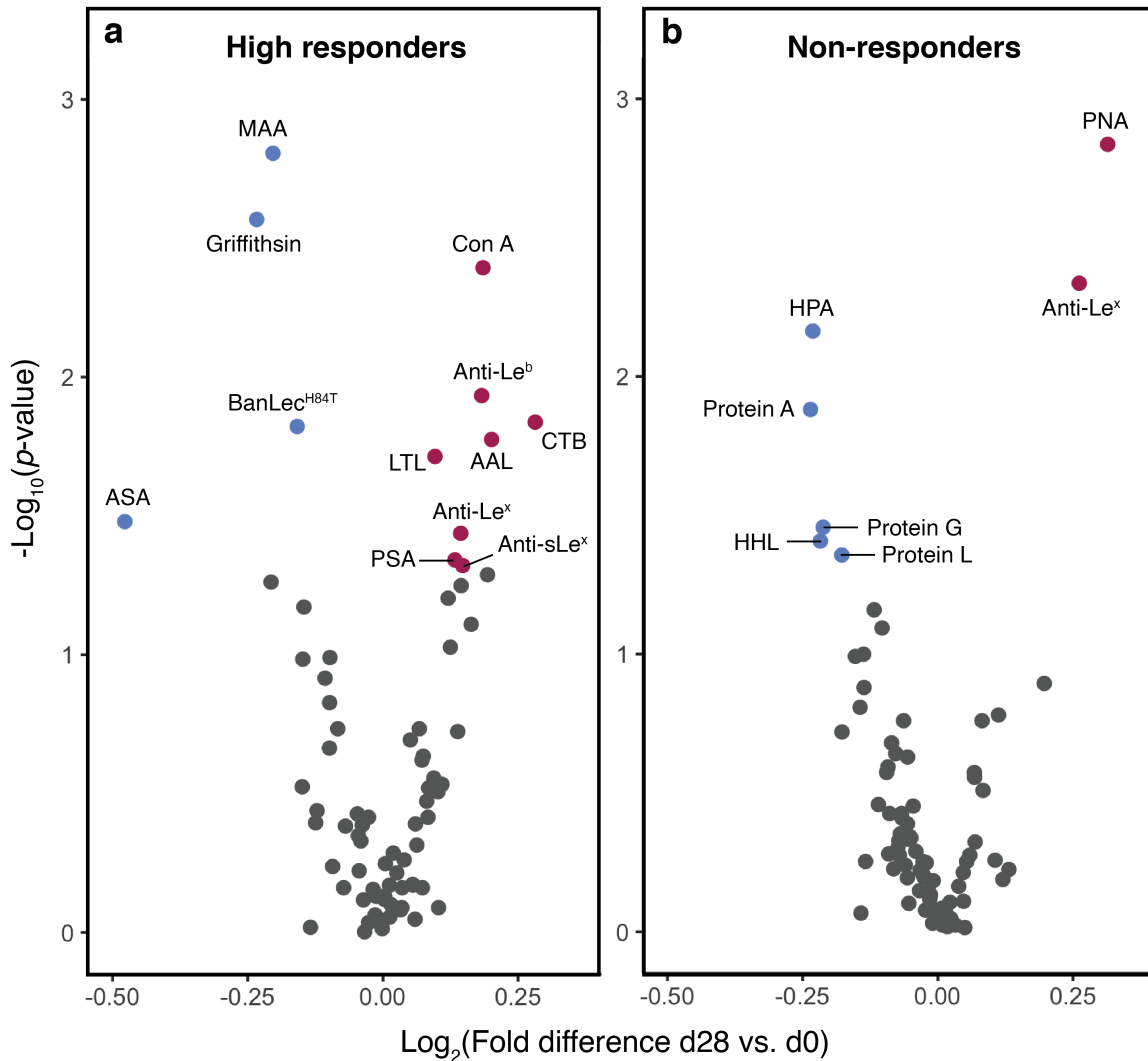


Figure 4.9. Volcano plot comparing the fold differences (d28 vs. d0) lectin microarray data for high responders (**a**; $n = 66$) and non-responders (**b**; $n = 54$). Mann–Whitney U test was used to determine p -values. Probes with $p < 0.05$ are labelled and colored in either maroon (upregulation) or blue (downregulation).

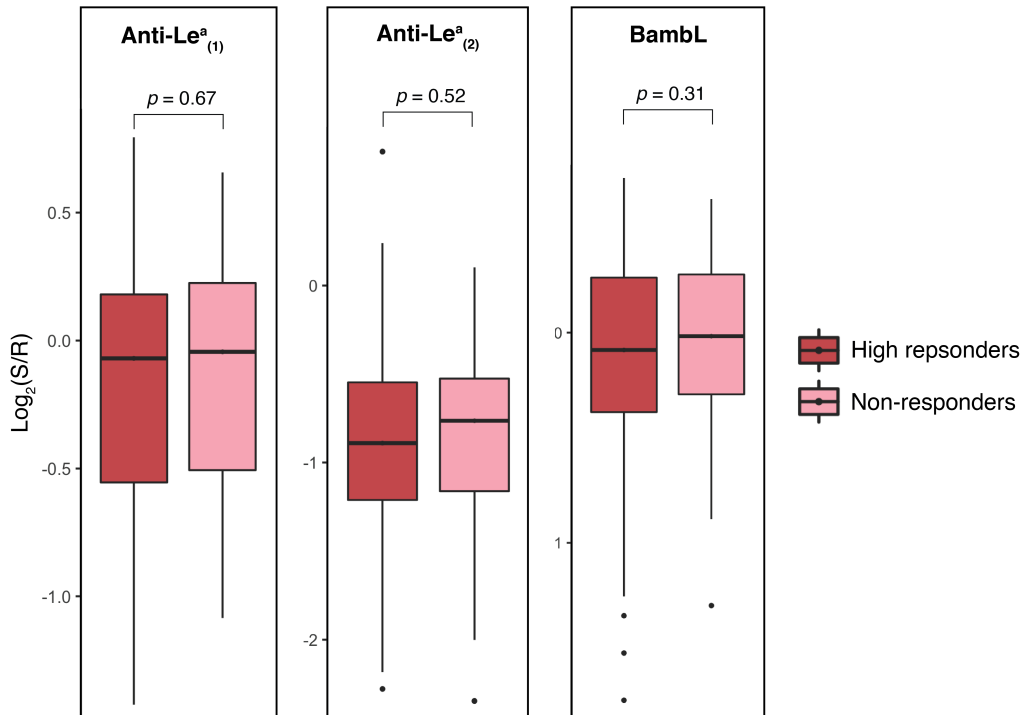


Figure 4.10. Boxplots comparing post-vaccination microarray data of Lewis A-binding lectins. Mann–Whitney U test was used to determine p -values.

4.3 Conclusions

The importance of vaccination has perhaps never been emphasized as much as today. For many vaccines, variability of vaccine response is a key issue that needs to be addressed.¹²⁶ Therefore, understanding the host factors influencing vaccine responses at the molecular level is essential. Host glycosylation, one of the pivotal players in immunity, has been surprisingly poorly studied in vaccination research. Data

from the present study argue that glycan signatures such as Le^a are associated with vaccine response, possibly through mechanisms involving the host microbiota. This finding could shed light on the development of novel immunomodulatory agents improving vaccine response, such as carbohydrate-based supplements.²¹⁹ In addition, serum Le^a was found highly enriched in proteins of the complement system. As roles of complement in effective vaccination are emerging, it would be beneficial to study whether glycosylation influences complement protein function, and therefore vaccine response. On the other hand, changes in glycosylation were observed in high responders, indicative of active immunomodulation. Future studies could look into why non-responders do not undergo these changes and whether these changes play a functional role in determining vaccine response.

4.4 Methods

Note: (1) This section (4.4) contains reprint from a publication. Reprinted with permission from *J. Proteome Res.* DOI: 10.1021/acs.jproteome.2c00251. Copyright 2022 American Chemical Society. (2) Most of this section (4.4) are the methods of the experiments performed by me. Methods of the experiments by collaborators (i.e., IHC staining) can be found in the corresponding reference (*J. Proteome Res.* DOI: 10.1021/acs.jproteome.2c00251)

Cohort recruitment

160 Caucasian adults were enrolled at the University of Georgia Clinical and Translational Research Unit (Athens, Georgia, USA) from September 2019 to February 2020. All volunteers were enrolled with written, informed consent. Participants were excluded if they already received the seasonal influenza vaccine. Other exclusion criteria included acute or chronic conditions that would put the participant at risk for an adverse reaction to the blood draw or the flu vaccine (*e.g.*, Guillain-Barré syndrome or allergies to egg products), or conditions that could skew the analysis (*e.g.*, recent flu symptoms or steroid injections/medications). All participants received a FLUZONE™ (Sanofi Pasteur, Lyon, France), seasonal inactivated influenza vaccine. Most received a quadrivalent, standard dose formulation made up of 15µg HA per strain of A/H1N1 (A/Brisbane/02/2018), A/H3N2 (A/Kansas/14/2017), B/Yamagata (B/Phuket/3073/2013) and B/Victoria (B/Colorado/6/2017-like strain).

Definition of antibody responses

Response scores for each strain of influenza were calculated based on the fold changes of antibody titers (d28 titer / d0 titer). For each strain, antibody response was scored in the following steps: i) calculate the initial score by taking the logarithmic (base 2) value of the titer fold change; ii) change the score to zero if the d28 antibody titer is lower or equal to 20, a conventional cut-off for effective protection;^{230,231} iii) change the score to 4 if the initial score is greater than 4 (*i.e.*, an over 16-fold increase in titer). This is to prevent the total response score (see below) from being biased towards one single strain. This strain-specific score was used to categorize the participants into three

response groups: high responders (score ≥ 2), low/moderate responders ($1 \leq \text{score} < 2$) and non-responders (score < 1) for each strain. The total response score is the sum of the four strain-specific scores. Similarly, total response scores were used to define overall high responders (score ≥ 8), overall low responders ($4 \leq \text{score} < 8$) and overall non-responders (score < 4). Note: people with low/moderate overall responses may be classified as high or non-responders in a strain-specific manner dependent on their response.

We also categorized antibody response using the metric of Wu *et al.* which takes into account BMI, age and gender.²⁰³ For our comparison, we ranked participants by the modified response score from high to low. The upper third was considered high-responders and the lower third non responders.

Lectin microarray

Total protein concentrations of serum samples were measured with *DC*TM protein assay kit (Bio-Rad Laboratories, Hercules, California, USA). Each volunteer serum sample was fluorescently labelled with Alexa Fluor 555TM NHS ester (Thermo Fisher Scientific, Waltham, Massachusetts, USA). First, 10 μg of total protein was diluted in PBS to 27 μL . The pH of the solution was adjusted with 3 μL of 1M sodium bicarbonate. Then 0.21 μL of a stock solution (10 mg/mL) of Alexa FluorTM 555 NHS ester was added to the mixture. The reaction lasted for 1 hour in the dark at room temperature. Unconjugated dye molecules were then removed by ZebaTM Dye and Biotin Removal Filter Plates (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The reference

material, NIST human serum 909c (Millipore Sigma, Darmstadt, Germany), was fluorescently labelled with Alexa Fluor 647™ NHS ester (Thermo Fisher Scientific, Waltham, Massachusetts, USA) similarly. The amounts of reagents were scaled linearly to the starting protein amount (4 mg). Finally, each Alexa Fluor 555-labelled sample (10 µg of total protein) was mixed with a proper volume of Alexa Fluor™ 647-labelled reference material containing the same amount of protein, and the final volume was adjusted to 50 µL with PBS.

Lectin microarray slides were fabricated as previously described.¹²⁷ The print was quality controlled as previously described.¹²⁷ Prior to hybridization, each dual-color mixture was diluted with 50 µL 0.2% PBST (PBS with 0.2% Tween-20, v/v). Each mixture was then allowed to hybridize with the arrays for 1 hour in the dark at room temperature. Arrays were washed twice with 0.005% PBST for 5 minutes and once with PBS for 5 minutes. The slides were briefly rinsed with ultrapure water and dried. Fluorescence signals were obtained with Genepix™ 4400A fluorescence slide scanner (Molecular Devices, San Jose, California, USA) in the 532 nm channel and the 635 nm channel that correspond to the excitation/emission profiles of Alexa Fluor™ 555 and Alexa Fluor™ 647, respectively. Raw fluorescence signal and background signal of each spot were generated by the Genepix Pro™ 7 software (Molecular Devices, San Jose, California, USA), which were further processed and analyzed with a custom script as previously described.⁸⁸ Heatmaps and volcano plots were generated with R (R version 4.0.1, r-project.org). Lectin annotation was done using data from Bojar et al.⁸⁶ In

general, epitopes are annotated when unambiguous (e.g., multiple related binders trending together).

Lectin/Antibody Affinity Pulldown

80 µg of BambL (expressed in-house) or anti-Le^a (Abcam, Cambridge, United Kingdom) was immobilized on columns using AminoLink™ Plus Micro Immobilization Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA) as per manufacturer's protocol. Coupling was carried out at 4°C overnight with gentle agitation. For the beads-only controls, PBS was added to the columns instead of BambL/anti-Le^a in the coupling step.

For protein identification by proteomics, a serum pool was prepared by combining 10 µL of each day 0 serum sample. The pooled serum was incubated at 54°C for 1 hour to inactivate proteases prior to the pulldown experiments. For BambL pulldown, 10 µL of pooled serum was diluted in PBS to 200 µL and incubated on the column for 1 hour at room temperature with gentle agitation. The column was washed with 300 µL PBS three times (5 minute per wash with gentle agitation). The column was eluted with 200 µL 50mM methyl α-L-fucopyranoside (TCI America, Portland, Oregon, USA) in PBS. For anti-Le^a pulldown, 50 µL of pooled serum was diluted in PBS to 400 µL and incubated on the column for 1 hour at room temperature with gentle agitation. The column was washed with 300 µL PBS three times (5 minute per wash with gentle agitation) before being eluted with 100 µL 0.1M glycine (pH = 2.8). The eluate was immediately neutralized with 30 µL 0.5M Tris (pH = 8.5). The protocol for preparing the six BambL-

pulldown samples for western blotting is the same as the protocol for glycoproteomics except that the columns were incubated with 200 μg serum protein diluted in 100 μL PBS, washed with 100 μL PBS and eluted with 80 μL elution buffer.

Mass spectrometry and Data Analysis

The enriched samples were incubated at 95°C for 10 min. 1 $\mu\text{g}/\mu\text{L}$ of sequencing grade-modified trypsin (Promega, Madison, Wisconsin, USA) was added to samples and overnight at 37°C with gentle agitation. Digestion was quenched by pH <4.0 using 2.5% trifluoroacetic acid (TFA). Samples were subsequently desalted using Pierce C18 spin tips (Thermo Fisher Scientific, Waltham, Massachusetts, USA) as per manufacturer's protocol. The peptides were eluted using aqueous buffer with 60% acetonitrile (ACN) and 0.1% formic acid (FA). The samples were dried, and peptides resuspended in 10 μl of buffer (0.1% FA in 5% ACN).

Each sample (~3 μL) was loaded onto Acclaim PepMap 100 trap column (75 μm x 2 cm) nanoViper, attached to an EASY-spray analytical column (PepMap RSLC C18, 2 μm , 100Å, 75 μm ID x 50 cm) in an EASY nano-LC 1000 liquid chromatography instrument (Thermo Scientific). Chromatography solvent A consisted of LC-MS grade water with 0.1% FA, and solvent B of 80% acetonitrile with 0.1% FA. The 155 min gradient consisted of: 2-5% of solvent B for 5 min, 5-25% for 110 min, 25-40% for 25 min, 40-80% for 5 min, 80-95% for 5 min, followed by 95-5% for 5 min. Mass spectrometry data was collected in data dependent mode on an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The MS1

spectra were recorded with a resolution of 240,000, AGC target of $1e6$, with maximum injection time of 50 ms, and a scan range of 400 to 1500 m/z. The MS2 spectra were collected using quadrupole isolation mode, AGC target of $2e4$, maximum injection time of 18 ms, one microscan, 0.7m/z isolation window, collision energy of 27%, excluding ions of charge state $<+2$ and $>+7$.

Spectra were searched against the Uniprot human fasta sequence database (UP000005640, downloaded on July 24, 2020) using the MaxQuant software (version 1.5.5.1) with default settings, including 2 missed cleavages, first search with peptide tolerance of 20 ppm and for the main search with peptide tolerance of 4.5 ppm. Carbamidomethylation of Cysteine was set as a static modification. The false discovery rates for peptide and protein identifications were both set to 0.01. Oxidation of Met and acetylation of the protein N terminus were the allowed variable modifications, and proteins were quantified using the Label Free Quantification (LFQ) option. A protein was identified as a positive binder if the enriched sample (E) and the corresponding control sample (C) satisfied the following: i) the sum of \log_{10} -transformed LFQ intensities of this protein in E and C was > 3 ; and ii) the difference of \log_2 -transformed LFQ intensities of this protein between E and C was > 2 (E-C). The remaining proteins were searched in GlyGen,²³² a database that compiles the experimental evidence for glycosylation of proteins. Proteins without any experimental evidence for glycosylation or solely with experimental evidence for *O*-GlcNAcylation were removed, as they are not the targets of interest of the pulldown experiments.

Pathway enrichment analysis

Gene Ontology (GO) enrichment analysis was performed with PANTHER Overrepresentation Test (Released 20210224).^{233,234} A full list of plasma proteins was used as the reference list.²⁸ “GO biological process complete”, “Fisher’s Exact” and “Calculate False Discovery Rate” were selected as the annotation data set, test type and correction method, respectively.

Western blotting

All steps were performed at room temperature. 20 µg of pulldown samples or input serum samples were resolved by 4-20% SDS/PAGE and transferred to a nitrocellulose membrane. Total protein was stained with Revert™ 700 Total Protein Stain Kit (LI-COR, Lincoln, Nebraska, USA) as per manufacturer’s instruction. After the total protein stain was erased, the membrane was blocked with blocking buffer [PBS with 3% (w/v) BSA and 0.05% (v/v) Tween-20] for 1 hour. Then the membrane was incubated with primary antibody working solution [rabbit anti-C4BPA (Abcam, Cambridge, United Kingdom), diluted to 0.5µg/ml in blocking buffer] for 1 hour, washed with PBST three times for 5 minutes per wash, incubated with secondary antibody working solution [goat anti-rabbit IgG CF™640R conjugate (Millipore Sigma, Darmstadt, Germany), diluted to 0.1µg/ml in blocking buffer] for 15 minutes and washed with PBST three times for 5 minutes per wash before imaging.

References

1. Varki, A. Biological roles of glycans. *Glycobiology* **27**, 3–49 (2017).
2. Bullard, W., Lopes Da Rosa-Spiegler, J., Liu, S., Wang, Y. & Sabatini, R. Identification of the glucosyltransferase that converts hydroxymethyluracil to base J in the trypanosomatid genome. *J. Biol. Chem.* **289**, 20273–20282 (2014).
3. Bowles, D., Lim, E. K., Poppenberger, B. & Vaistij, F. E. Glycosyltransferases of lipophilic small molecules. *Annu. Rev. Plant Biol.* **57**, 567–597 (2006).
4. A. Varki et al. *Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor Laboratory Press.* (Cold Spring Harbor Laboratory Press, 2017).
5. Rakus, J. F. & Mahal, L. K. New technologies for glycomic analysis: Toward a systematic understanding of the glycome. *Annu. Rev. Anal. Chem.* **4**, 367–392 (2011).
6. Loke, I., Østergaard, O., Heegaard, N. H. H., Packer, N. H. & Thaysen-Andersen, M. Paucimannose-rich N-glycosylation of spatiotemporally regulated Human neutrophil elastase modulates its immune functions. *Mol. Cell. Proteomics* **16**, 1507–1527 (2017).
7. Chen, S., Qin, R. & Mahal, L. K. Sweet systems: technologies for glycomic analysis and their integration into systems biology. *Crit. Rev. Biochem. Mol. Biol.* **56**, 301–320 (2021).
8. Mimura, Y. *et al.* Glycosylation engineering of therapeutic IgG antibodies: challenges for the safety, functionality and efficacy. *Protein Cell* **9**, 47–62 (2018).

9. Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: Diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **13**, 448–462 (2012).
10. Cummings, R. D. The repertoire of glycan determinants in the human glycome. *Mol. Biosyst.* **5**, 1087–1104 (2009).
11. Neelamegham, S. & Mahal, L. K. Multi-level regulation of cellular glycosylation: from genes to transcript to enzyme to structure. *Curr. Opin. Struct. Biol.* **40**, 145–152 (2016).
12. Ko, J. H. *et al.* Regulation of the GnT-V promoter by transcription factor Ets-1 in various cancer cell lines. *J. Biol. Chem.* **274**, 22941–22948 (1999).
13. Dorsett, K. A., Jones, R. B., Ankenbauer, K. E., Hjelmeland, A. B. & Bellis, S. L. Sox2 promotes expression of the ST6Gal-I glycosyltransferase in ovarian cancer cells. *J. Ovarian Res.* **12**, 1–13 (2019).
14. Wong, M. Y. *et al.* XBP1s activation can globally remodel N-glycan structure distribution patterns. *Proc. Natl. Acad. Sci.* **115**, E10089 LP-E10098 (2018).
15. Kroes, R. A. & Moskal, J. R. The role of DNA methylation in ST6Gal1 expression in gliomas. *Glycobiology* **26**, 1271–1283 (2016).
16. 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).
17. Vigetti, D. *et al.* Role of UDP-N-acetylglucosamine (GlcNAc) and O-GlcNAcylation of hyaluronan synthase 2 in the control of chondroitin sulfate and hyaluronan synthesis. *J. Biol. Chem.* **287**, 35544–35555 (2012).

18. Qin, R. & Mahal, L. K. The host glycomic response to pathogens. *Curr. Opin. Struct. Biol.* **68**, 149–156 (2021).
19. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555 (2015).
20. Sjöwall, C. *et al.* Altered glycosylation of complexed native IgG molecules is associated with disease activity of systemic lupus erythematosus. *Lupus* **24**, 569–581 (2015).
21. Sun, Q., Zhang, Z., Zhang, H. & Liu, X. Aberrant IgA1 glycosylation in iga nephropathy: A systematic review. *PLoS One* **11**, 1–15 (2016).
22. Demus, D. *et al.* Interlaboratory evaluation of plasma N-glycan antennary fucosylation as a clinical biomarker for HNF1A-MODY using liquid chromatography methods. *Glycoconj. J.* **38**, 375–386 (2021).
23. Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* **15**, 346–366 (2019).
24. Varki, A. *et al.* Symbol Nomenclature for Graphical Representations of Glycans. **25**, 1323–1324 (2015).
25. Punt, J., Stranford, S. A., Jones, P. P. & Owen, J. A. *Kuby immunology*. (Macmillan Education, 2019).
26. van Kooyk, Y. & Rabinovich, G. A. Protein-glycan interactions in the control of innate and adaptive immune responses. *Nat. Immunol.* **9**, 593–601 (2008).
27. Astronomo, R. D. & Burton, D. R. Carbohydrate vaccines: Developing sweet solutions to sticky situations? *Nat. Rev. Drug Discov.* **9**, 308–324 (2010).

28. Da Silva Correia, J. & Ulevitch, R. J. MD-2 and TLR4 N-linked glycosylations are important for a functional lipopolysaccharide receptor. *J. Biol. Chem.* **277**, 1845–1854 (2002).
29. Margraf-Schönfeld, S., Böhm, C. & Watzl, C. Glycosylation affects ligand binding and function of the activating natural killer cell receptor 2B4 (CD244) protein. *J. Biol. Chem.* **286**, 24142–24149 (2011).
30. Sun, J. *et al.* Structural and functional analyses of the human toll-like receptor 3: Role of glycosylation. *J. Biol. Chem.* **281**, 11144–11151 (2006).
31. Gerard, C., Chenoweth, D. E. & Hugli, T. E. Response of human neutrophils to C5a: a role for the oligosaccharide moiety of human C5ades Arg-74 but not of C5a in biologic activity. *J. Immunol.* **127**, 1978–1982 (1981).
32. Liszewski, M. K., Leung, M. K. & Atkinson, J. P. Membrane cofactor protein: Importance of N- and O-glycosylation for complement regulatory function. *Mol. Immunol.* **35**, 404 (1998).
33. Kontermann, R. & Rauterberg, E. W. N-Deglycosylation of human complement component C9 reduces its hemolytic activity. *Mol. Immunol.* **26**, 1125–1132 (1989).
34. Skořepa, O. *et al.* Natural killer cell activation receptor NKp30 oligomerization depends on its N-glycosylation. *Cancers (Basel)*. **12**, 1–24 (2020).
35. Mortales, C. L., Lee, S. U., Manousadjian, A., Hayama, K. L. & Demetriou, M. N-Glycan Branching Decouples B Cell Innate and Adaptive Immunity to Control Inflammatory Demyelination. *iScience* **23**, 101380 (2020).

36. Su, Y. *et al.* Glycosylation influences the lectin activities of the macrophage mannose receptor. *J. Biol. Chem.* **280**, 32811–32820 (2005).
37. Karmakar, J., Roy, S. & Mandal, C. Modulation of TLR4 Sialylation Mediated by a Sialidase Neu1 and Impairment of Its Signaling in *Leishmania donovani* Infected Macrophages. *Front. Immunol.* **10**, 1–19 (2019).
38. Feng, C. *et al.* Sialyl residues modulate LPS-mediated signaling through the toll-like receptor 4 complex. *PLoS One* **7**, (2012).
39. Blaum, B. S. *et al.* Structural basis for sialic acid-mediated self-recognition by complement factor H. *Nat. Chem. Biol.* **11**, 77–82 (2015).
40. Langford-Smith, A., Day, A. J., Bishop, P. N. & Clark, S. J. Complementing the sugar code: Role of GAGs and sialic acid in complement regulation. *Front. Immunol.* **6**, 1–7 (2015).
41. Glasner, A. *et al.* Identification of putative novel O-glycosylations in the NK killer receptor Ncr1 essential for its activity. *Cell Discov.* **1**, 1–14 (2015).
42. Merle, N. S., Noe, R., Halbwachs-Mecarelli, L., Fremeaux-Bacchi, V. & Roumenina, L. T. Complement system part II: Role in immunity. *Front. Immunol.* **6**, 1–26 (2015).
43. Merle, N. S., Church, S. E., Fremeaux-Bacchi, V. & Roumenina, L. T. Complement system part I - molecular mechanisms of activation and regulation. *Front. Immunol.* **6**, 1–30 (2015).
44. Kuipers, S., Aerts, P. C. & Van Dijk, H. Differential microorganism-induced mannose-binding lectin activation. *FEMS Immunol. Med. Microbiol.* **36**, 33–39

- (2003).
45. Ritchie, G. E. *et al.* Glycosylation and the complement system. *Chem. Rev.* **102**, 305–319 (2002).
 46. Irvine, E. B. & Alter, G. Understanding the role of antibody glycosylation through the lens of severe viral and bacterial diseases. *Glycobiology* **30**, 241–253 (2020).
 47. Mi, Y. *et al.* Functional consequences of mannose and asialoglycoprotein receptor ablation. *J. Biol. Chem.* **291**, 18700–18717 (2016).
 48. Bas, M. *et al.* Fc Sialylation Prolongs Serum Half-Life of Therapeutic Antibodies. *J. Immunol.* **202**, 1582–1594 (2019).
 49. Goetze, A. M. *et al.* High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology* **21**, 949–959 (2011).
 50. 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).
 51. Steffen, U. *et al.* IgA subclasses have different effector functions associated with distinct glycosylation profiles. *Nat. Commun.* **11**, (2020).
 52. Shade, K. T. C. *et al.* Sialylation of immunoglobulin E is a determinant of allergic pathogenicity. *Nature* **582**, 265–270 (2020).
 53. Colucci, M. *et al.* Sialylation of N-Linked Glycans Influences the Immunomodulatory Effects of IgM on T Cells. *J. Immunol.* **194**, 151–157 (2015).
 54. Giannone, C., Fagioli, C., Valetti, C., Sitia, R. & Anelli, T. Roles of N-glycans in

- the polymerization-dependent aggregation of mutant Ig- μ chains in the early secretory pathway. *Sci. Rep.* **7**, 1–10 (2017).
55. Sun, R., Kim, A. M. J. & Lim, S. O. Glycosylation of Immune Receptors in Cancer. *Cells* **10**, (2021).
 56. Okada, M. *et al.* Blockage of Core Fucosylation Reduces Cell-Surface Expression of PD-1 and Promotes Anti-tumor Immune Responses of T Cells. *Cell Rep.* **20**, 1017–1028 (2017).
 57. Edgar, L. J. *et al.* Sialic Acid Ligands of CD28 Suppress Costimulation of T Cells. *ACS Cent. Sci.* **7**, 1508–1515 (2021).
 58. Johannes, L., Jacob, R. & Leffler, H. Galectins at a glance. *J. Cell Sci.* **131**, 1–9 (2018).
 59. Movsisyan, L. D. & Macauley, M. S. Structural advances of Siglecs: Insight into synthetic glycan ligands for immunomodulation. *Org. Biomol. Chem.* **18**, 5784–5797 (2020).
 60. Tvaroška, I., Selvaraj, C. & Koča, J. Selectins—The Two Dr. Jekyll and Mr. Hyde Faces of Adhesion Molecules—A Review. *Molecules* **25**, 2835 (2020).
 61. Poe, J. C. & Tedder, T. F. CD22 and Siglec-G in B cell function and tolerance. *Trends Immunol.* **33**, 413–420 (2012).
 62. Ludwig, A. K. *et al.* Design–functionality relationships for adhesion/ growth-regulatory galectins. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 2837–2842 (2019).
 63. Demetriou, M., Granovsky, M., Quaggin, S. & Dennis, J. W. Tissue triglyceride content Adipocyte TNF- α mRNA and serum TNF- α levels. **409**, 733–739 (2001).

64. van Houtum, E. J. H., Büll, C., Cornelissen, L. A. M. & Adema, G. J. Siglec Signaling in the Tumor Microenvironment. *Front. Immunol.* **12**, 1–17 (2021).
65. Li, F. & Ding, J. Sialylation is involved in cell fate decision during development, reprogramming and cancer progression. *Protein Cell* **10**, 550–565 (2019).
66. Mantuano, N. R., Natoli, M., Zippelius, A. & Läubli, H. Tumor-associated carbohydrates and immunomodulatory lectins as targets for cancer immunotherapy. *J. Immunother. Cancer* **8**, 1–12 (2020).
67. Alter, G. *et al.* High-resolution definition of humoral immune response correlates of effective immunity against HIV . *Mol. Syst. Biol.* **14**, 1–14 (2018).
68. Lu, L. L. *et al.* A Functional Role for Antibodies in Tuberculosis. *Cell* **167**, 433-443.e14 (2016).
69. Vadrevu, S. K. *et al.* Frontline Science: Plasma and immunoglobulin G galactosylation associate with HIV persistence during antiretroviral therapy. *J. Leukoc. Biol.* **104**, 461–471 (2018).
70. Chakraborty, S. *et al.* Proinflammatory IgG Fc structures in patients with severe COVID-19. *Nat. Immunol.* **22**, 67–73 (2021).
71. Larsen, M. D. *et al.* Afucosylated IgG characterizes enveloped viral responses and correlates with COVID-19 severity. *Science (80-.)*. **371**, (2021).
72. Lu, L. L. *et al.* Antibody Fc Glycosylation Discriminates Between Latent and Active Tuberculosis. *J. Infect. Dis.* **02139**, 1–10 (2020).
73. Naegeli, A. *et al.* Streptococcus pyogenes evades adaptive immunity through specific IgG glycan hydrolysis. *J. Exp. Med.* **216**, 1615–1629 (2019).

74. Heindel, D. W. *et al.* Glycomic analysis of host response reveals high mannose as a key mediator of influenza severity. *Proc. Natl. Acad. Sci.* **117**, 26926–26935 (2020).
75. Libertucci, J. & Young, V. B. The role of the microbiota in infectious diseases. *Nat. Microbiol.* **4**, 35–45 (2019).
76. Pickard, J. M. *et al.* Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness. *Nature* **514**, 638–641 (2014).
77. Pham, T. A. N. *et al.* Epithelial IL-22RA1-mediated fucosylation promotes intestinal colonization resistance to an opportunistic pathogen. *Cell Host Microbe* **16**, 504–516 (2014).
78. Nagao-Kitamoto, H. *et al.* Interleukin-22-mediated host glycosylation prevents *Clostridioides difficile* infection by modulating the metabolic activity of the gut microbiota. *Nat. Med.* **26**, 608–617 (2020).
79. Theodoratou, E. *et al.* The role of glycosylation in IBD. *Nat. Rev. Gastroenterol. Hepatol.* **11**, 588–600 (2014).
80. Tirumalai, R. S. *et al.* Characterization of the low molecular weight human serum proteome. *Mol. Cell. Proteomics* **2**, 1096–1103 (2003).
81. Nitschke, L. CD22 and Siglec-G regulate inhibition of B-cell signaling by sialic acid ligand binding and control B-cell tolerance. *Glycobiology* **24**, 807–817 (2014).
82. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat. Rev. Immunol.* **16**, 626–638 (2016).
83. Forsey, R. J. *et al.* Plasma cytokine profiles in elderly humans. *Mech. Ageing Dev.*

- 124**, 487–493 (2003).
84. Michaud, M. *et al.* Proinflammatory cytokines, aging, and age-related diseases. *J. Am. Med. Dir. Assoc.* **14**, 877–882 (2013).
 85. Dotz, V. & Wuhrer, M. N-glycome signatures in human plasma: associations with physiology and major diseases. *FEBS Lett.* **593**, 2966–2976 (2019).
 86. Bojar, D. *et al.* A Useful Guide to Lectin Binding: Machine-Learning Directed Annotation of 57 Unique Lectin Specificities. *ACS Chem. Biol.* (2022)
doi:10.1021/acscchembio.1c00689.
 87. Pilobello, K. T., Slawek, D. E. & Mahal, L. K. A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 11534–11539 (2007).
 88. Batista, B. S., Eng, W. S., Pilobello, K. T., Hendricks-Muñoz, K. D. & Mahal, L. K. Identification of a conserved glycan signature for microvesicles. *J. Proteome Res.* **10**, 4624–4633 (2011).
 89. Chen, S. *et al.* Age-Dependent Glycomic Response to the 2009 Pandemic H1N1 Influenza Virus and Its Association with Disease Severity. *J. Proteome Res.* **19**, 4486–4495 (2020).
 90. Daley, D. *et al.* Dectin 1 activation on macrophages by galectin 9 promotes pancreatic carcinoma and peritumoral immune tolerance. *Nat. Med.* **23**, 556–567 (2017).
 91. Kurz, E. *et al.* Integrated Systems Analysis of the Murine and Human Pancreatic Cancer Glycomes Reveals a Tumor-Promoting Role for ST6GAL1. *Mol. Cell.*

- Proteomics* **20**, 100160 (2021).
92. Agrawal, P. *et al.* A Systems Biology Approach Identifies FUT8 as a Driver of Melanoma Metastasis. *Cancer Cell* **31**, 804-819.e7 (2017).
 93. Grigorian, A. & Demetriou, M. Mgat5 Deficiency in T Cells and Experimental Autoimmune Encephalomyelitis . *ISRN Neurol.* **2011**, 1–6 (2011).
 94. Demetriou, M., Granovsky, M., Quaggin, S. & Dennis, J. W. Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation. *Nature* **409**, 733–739 (2001).
 95. Cabral, J. *et al.* Distinctive surface glycosylation patterns associated with mouse and human cD4+ regulatory T cells and their suppressive function. *Front. Immunol.* **8**, (2017).
 96. Morgan, R. *et al.* N -Acetylglucosaminyltransferase V (Mgat5)-Mediated N -Glycosylation Negatively Regulates Th1 Cytokine Production by T Cells . *J. Immunol.* **173**, 7200–7208 (2004).
 97. Huffman, J. E. *et al.* Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults. *Hum. Mol. Genet.* **20**, 5000–5011 (2011).
 98. Pereira, M. S. *et al.* Genetic Variants of the MGAT5 Gene Are Functionally Implicated in the Modulation of T Cells Glycosylation and Plasma IgG Glycome Composition in Ulcerative Colitis. *Clin. Transl. Gastroenterol.* **11**, e00166 (2020).
 99. Cooling, L. Blood groups in infection and host susceptibility. *Clin. Microbiol. Rev.* **28**, 801–870 (2015).

100. Matsumoto, Y. *et al.* Identification of Tn antigen O-GalNAc-expressing glycoproteins in human carcinomas using novel anti-Tn recombinant antibodies. *Glycobiology* **30**, 282–300 (2020).
101. Hoffmann, M., Marx, K., Reichl, U., Wuhrer, M. & Rapp, E. Site-specific O-glycosylation analysis of human blood plasma proteins. *Mol. Cell. Proteomics* **15**, 624–641 (2016).
102. Hoiland, R. L. *et al.* The association of ABO blood group with indices of disease severity and multiorgan dysfunction in COVID-19. *Blood Adv.* **4**, 4981–4989 (2020).
103. Bensing, B. A., Li, Q., Park, D., Lebrilla, C. B. & Sullam, P. M. Streptococcal Siglec-like adhesins recognize different subsets of human plasma glycoproteins: Implications for infective endocarditis. *Glycobiology* **28**, 601–611 (2018).
104. Bergstrom, K., Fu, J. & Xia, L. Biological Functions of C1GalT1 and Mucin-Type O-Glycans BT - Glycoscience: Biology and Medicine. in (eds. Taniguchi, N., Endo, T., Hart, G. W., Seeberger, P. H. & Wong, C.-H.) 1073–1080 (Springer Japan, 2015). doi:10.1007/978-4-431-54841-6_65.
105. Lee, P. C. *et al.* C1GALT1 is associated with poor survival and promotes soluble Ephrin A1-mediated cell migration through activation of EPHA2 in gastric cancer. *Oncogene* **39**, 2724–2740 (2020).
106. Steirer, L. M., Park, E. I., Townsend, R. R. & Baenzinger, J. U. The asialoglycoprotein receptor regulates levels of plasma glycoproteins terminating with sialic acid α 2,6-galactose. *J. Biol. Chem.* **284**, 3777–3783 (2009).

107. Park, E. I., Manzella, S. M. & Baenziger, J. U. Rapid clearance of sialylated glycoproteins by the asialoglycoprotein receptor. *J. Biol. Chem.* **278**, 4597–4602 (2003).
108. Ribeiro, J. P. *et al.* Characterization of a high-affinity sialic acid-specific CBM40 from *Clostridium perfringens* and engineering of a divalent form. *Biochem. J.* **473**, 2109–2118 (2016).
109. Englyst, N. A. *et al.* Percentage of body fat and plasma glucose predict plasma sialic acid concentration in type 2 diabetes mellitus. *Metabolism.* **55**, 1165–1170 (2006).
110. Nielsen, M. J. *et al.* Haptoglobin-related protein is a high-affinity hemoglobin-binding plasma protein. *Blood* **108**, 2846–2849 (2006).
111. Chiellini, C. *et al.* Obesity modulates the expression of haptoglobin in the white adipose tissue via TNF α . *J. Cell. Physiol.* **190**, 251–258 (2002).
112. Chiellini, C. *et al.* Serum haptoglobin: A novel marker of adiposity in humans. *J. Clin. Endocrinol. Metab.* **89**, 2678–2683 (2004).
113. Vázquez-Moreno, M. *et al.* Causal association of haptoglobin with obesity in mexican children: A mendelian randomization study. *J. Clin. Endocrinol. Metab.* **105**, 2501–2510 (2020).
114. Oller Moreno, S. *et al.* The differential plasma proteome of obese and overweight individuals undergoing a nutritional weight loss and maintenance intervention. *Proteomics - Clin. Appl.* **12**, 1–22 (2018).
115. Novince, C. M. *et al.* Proteoglycan 4, a novel immunomodulatory factor, regulates

- parathyroid hormone actions on hematopoietic cells. *Am. J. Pathol.* **179**, 2431–2442 (2011).
116. Iqbal, S. M. *et al.* Lubricin/Proteoglycan 4 binds to and regulates the activity of Toll-Like Receptors in Vitro. *Sci. Rep.* **6**, 1–12 (2016).
117. Ali, L. *et al.* The O-glycomap of lubricin, a novel mucin responsible for joint lubrication, identified by site-specific glycopeptide analysis. *Mol. Cell. Proteomics* **13**, 3396–3409 (2014).
118. Zaytseva, O. O. *et al.* Heritability of Human Plasma N-Glycome. *J. Proteome Res.* **19**, 85–91 (2020).
119. Knežević, A. *et al.* Variability, heritability and environmental determinants of human plasma n-glycome. *J. Proteome Res.* **8**, 694–701 (2009).
120. 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).
121. Ding, N. *et al.* Human serum N-glycan profiles are age and sex dependent. *Age Ageing* **40**, 568–575 (2011).
122. Tijardović, M. *et al.* Intense Physical Exercise Induces an Anti-inflammatory Change in IgG N-Glycosylation Profile. *Front. Physiol.* **10**, 1–10 (2019).
123. Štambuk, J. *et al.* Global variability of the human IgG glycome. *Aging (Albany, NY)*. **12**, 1–13 (2020).
124. Andrade, F. B. *et al.* The Weight of Obesity in Immunity from Influenza to COVID-19. *Front. Cell. Infect. Microbiol.* **11**, 1–14 (2021).
125. Mclaughlin, T., Ackerman, S. E., Shen, L. & Engleman, E. Role of innate and

- adaptive immunity in obesity-associated metabolic disease. *J. Clin. Invest.* **127**, 5–13 (2017).
126. Dhakal, S. & Klein, S. L. Host Factors Impact Vaccine Efficacy: Implications for Seasonal and Universal Influenza Vaccine Programs. *J. Virol.* 1–15 (2019).
127. Pilobello, K. T., Agrawal, P., Rouse, R. & Mahal, L. K. Advances in lectin microarray technology: optimized protocols for piezoelectric print conditions. *Curr. Protoc. Chem. Biol.* **5**, 1–23 (2013).
128. World Health Organization. WHO COVID-19 Dashboard. <https://covid19.who.int/> (2020).
129. Hu, B., Guo, H., Zhou, P. & Shi, Z.-L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **19**, 141–154 (2021).
130. Romagnoli, S., Peris, A., De Gaudio, A. R. & Geppetti, P. SARS-CoV-2 and COVID-19: From the bench to the bedside. *Physiol. Rev.* **100**, 1455–1466 (2020).
131. Alosaimi, B. *et al.* Complement Anaphylatoxins and Inflammatory Cytokines as Prognostic Markers for COVID-19 Severity and In-Hospital Mortality . *Frontiers in Immunology* vol. 12 (2021).
132. Qin, G., Liu, S., Yang, L., Yu, W. & Zhang, Y. Myeloid cells in COVID-19 microenvironment. *Signal Transduct. Target. Ther.* **6**, 372 (2021).
133. Zuo, Y. *et al.* Autoantibodies stabilize neutrophil extracellular traps in COVID-19. *JCI Insight* **6**, (2021).
134. Wang, E. Y. *et al.* Diverse functional autoantibodies in patients with COVID-19. *Nature* **595**, 283–288 (2021).

135. Afzali, B., Noris, M., Lambrecht, B. N. & Kemper, C. The state of complement in COVID-19. *Nat. Rev. Immunol.* **22**, 77–84 (2022).
136. Petrović, T. *et al.* Composition of the immunoglobulin G glycome associates with the severity of COVID-19. *Glycobiology* **31**, 372–377 (2021).
137. Röltgen, K. *et al.* Defining the features and duration of antibody responses to SARS-CoV-2 infection associated with disease severity and outcome. *Sci. Immunol.* **5**, 1–20 (2020).
138. Kollias, A. *et al.* Thromboembolic risk and anticoagulant therapy in COVID-19 patients: emerging evidence and call for action. *Br. J. Haematol.* **189**, 846–847 (2020).
139. Zou, F., Qian, Z., Wang, Y., Zhao, Y. & Bai, J. Cardiac Injury and COVID-19: A Systematic Review and Meta-analysis. *CJC Open* **2**, 386–394 (2020).
140. Legrand, M. *et al.* Pathophysiology of COVID-19-associated acute kidney injury. *Nat. Rev. Nephrol.* **17**, 751–764 (2021).
141. Xie, Y., Xu, E., Bowe, B. & Al-Aly, Z. Long-term cardiovascular outcomes of COVID-19. *Nat. Med.* **28**, 583–590 (2022).
142. Koo, H. J. *et al.* Radiographic and CT Features of Viral Pneumonia. *RadioGraphics* **38**, 719–739 (2018).
143. Haseli, S., Khalili, N., Bakhshayeshkaram, M., Sanei-Taheri, M. & Moharramzad, Y. Lobar Distribution of COVID-19 Pneumonia Based on Chest Computed Tomography Findings; A Retrospective Study. *Arch. Acad. Emerg. Med.* **8**, e55 (2020).

144. Shi, H. *et al.* Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect. Dis.* **20**, 425–434 (2020).
145. Bolognesi, B. & Lehner, B. Protein Overexpression: Reaching the limit. *Elife* **7**, e39804 (2018).
146. Kitagawa, H. & Paulson, J. C. Differential expression of five sialyltransferase genes in human tissues. *J. Biol. Chem.* **269**, 17872–17878 (1994).
147. O'Hanlon, T. P., Lau, K. M., Wang, X. C. & Lau, J. T. Tissue-specific Expression of β -Galactoside α -2,6-Sialyltransferase. *J. Biol. Chem.* **264**, 17389–17394 (1989).
148. Shen, B. *et al.* Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59-72.e15 (2020).
149. Wang, X. *et al.* Dysregulated hematopoiesis in bone marrow marks severe COVID-19. *Cell Discov.* **7**, 1–18 (2021).
150. Margaroli, C. *et al.* Spatial mapping of SARS-CoV-2 and H1N1 lung injury identifies differential transcriptional signatures. *Cell Reports Med.* **2**, 100242 (2021).
151. Rapkiewicz, A. V. *et al.* Megakaryocytes and platelet-fibrin thrombi characterize multi-organ thrombosis at autopsy in COVID-19: A case series. *EClinicalMedicine* **24**, (2020).
152. Wool, G. D. & Miller, J. L. The Impact of COVID-19 Disease on Platelets and Coagulation. *Pathobiology* **88**, 15–27 (2021).
153. Di Micco, P. *et al.* Prognostic value of fibrinogen among covid-19 patients

- admitted to an emergency department: An Italian cohort study. *J. Clin. Med.* **9**, 1–6 (2020).
154. Sui, J., Noubouossie, D. F., Gandotra, S. & Cao, L. Elevated Plasma Fibrinogen Is Associated With Excessive Inflammation and Disease Severity in COVID-19 Patients. *Front. Cell. Infect. Microbiol.* **11**, 1–7 (2021).
155. Lo, M. W., Kemper, C. & Woodruff, T. M. COVID-19: Complement, Coagulation, and Collateral Damage. *J. Immunol.* **205**, 1488–1495 (2020).
156. Noris, M., Benigni, A. & Remuzzi, G. The case of complement activation in COVID-19 multiorgan impact. *Kidney Int.* **98**, 314–322 (2020).
157. Zuo, Y. *et al.* Neutrophil extracellular traps (NETs) as markers of disease severity in COVID-19. *medRxiv* 2020.04.09.20059626 (2020)
doi:10.1101/2020.04.09.20059626.
158. Kulkarni, H. S. & Atkinson, J. P. Targeting complement activation in COVID-19. *Blood* **136**, 2000–2001 (2020).
159. Stahel, P. F. & Barnum, S. R. Complement Inhibition in Coronavirus Disease (COVID)-19: A Neglected Therapeutic Option . *Frontiers in Immunology* vol. 11 (2020).
160. Risitano, A. M. *et al.* Complement as a target in COVID-19? *Nat. Rev. Immunol.* **20**, 343–344 (2020).
161. Fang, C., Zhang, X., Miwa, T. & Song, W.-C. Complement promotes the development of inflammatory T-helper 17 cells through synergistic interaction with Toll-like receptor signaling and interleukin-6 production. *Blood* **114**, 1005–1015

- (2009).
162. Katz, Y., Revel, M. & Strunk, R. C. Interleukin 6 stimulates synthesis of complement proteins factor B and C3 in human skin fibroblasts. *Eur. J. Immunol.* **19**, 983–988 (1989).
 163. Kermali, M., Khalsa, R. K., Pillai, K., Ismail, Z. & Harky, A. The role of biomarkers in diagnosis of COVID-19 – A systematic review. *Life Sci* **254**, 117788 (2020).
 164. Liu, Y. *et al.* Complement C3 produced by macrophages promotes renal fibrosis via IL-17A secretion. *Front. Immunol.* **9**, 1–17 (2018).
 165. Sandler, M. *et al.* Complement Component 5 Mediates Development of Fibrosis, via Activation of Stellate Cells, in 2 Mouse Models of Chronic Pancreatitis. *Gastroenterology* **149**, 765-776.e10 (2015).
 166. Kou, W. *et al.* High complement protein C1q levels in pulmonary fibrosis and non-small cell lung cancer associated with poor prognosis. *BMC Cancer* **22**, 110 (2022).
 167. Okamoto, T. *et al.* The relationship between complement C3 expression and the MUC5B genotype in pulmonary fibrosis. *Am. J. Physiol. Cell. Mol. Physiol.* **315**, L1–L10 (2018).
 168. Addis-Lieser, E., Köhl, J. & Chiaramonte, M. G. Opposing Regulatory Roles of Complement Factor 5 in the Development of Bleomycin-Induced Pulmonary Fibrosis. *J. Immunol.* **175**, 1894–1902 (2005).
 169. Hair, P. S. *et al.* Complement effectors, C5a and C3a, in cystic fibrosis lung fluid correlate with disease severity. *PLoS One* **12**, e0173257 (2017).

170. Wendisch, D. *et al.* SARS-CoV-2 infection triggers profibrotic macrophage responses and lung fibrosis. *Cell* **184**, 6243-6261.e27 (2021).
171. D'Alessandro, A. *et al.* Serum Proteomics in COVID-19 Patients: Altered Coagulation and Complement Status as a Function of IL-6 Level. *J. Proteome Res.* **19**, 4417–4427 (2020).
172. Whetton, A. D., Preston, G. W., Abubeker, S. & Geifman, N. Proteomics and Informatics for Understanding Phases and Identifying Biomarkers in COVID-19 Disease. *J. Proteome Res.* **19**, 4219–4232 (2020).
173. Zelek, W. M. *et al.* Complement Inhibition with the C5 Blocker LFG316 in Severe COVID-19. *Am. J. Respir. Crit. Care Med.* **202**, 1304–1308 (2020).
174. Xu, B. *et al.* SOD1 is a Possible Predictor of COVID-19 Progression as Revealed by Plasma Proteomics. *ACS Omega* **6**, 16826–16836 (2021).
175. Garner, B. *et al.* Structural elucidation of the N- and O-glycans of human apolipoprotein(a). Role of O-glycans in conferring protease resistance. *J. Biol. Chem.* **276**, 22200–22208 (2001).
176. Maher, J. A. & DeStefano, J. The Ferret: An Animal Model to Study Influenza Virus. *Lab Anim. (NY)*. **33**, 50–53 (2004).
177. Shou, S. *et al.* Animal Models for COVID-19: Hamsters, Mouse, Ferret, Mink, Tree Shrew, and Non-human Primates . *Frontiers in Microbiology* vol. 12 (2021).
178. Dong, A. *et al.* Protection of K18-hACE2 mice and ferrets against SARS-CoV-2 challenge by a single-dose mucosal immunization with a parainfluenza virus 5–

- based COVID-19 vaccine. *Sci. Adv.* **7**, eabi5246 (2022).
179. Francis, M. E. *et al.* Sex and age bias viral burden and interferon responses during SARS-CoV-2 infection in ferrets. *Sci. Rep.* **11**, 1–18 (2021).
 180. Jones, S. A. & Hunter, C. A. Is IL-6 a key cytokine target for therapy in COVID-19? *Nat. Rev. Immunol.* **21**, 337–339 (2021).
 181. Tamara, A. & Tahapary, D. L. Obesity as a predictor for a poor prognosis of COVID-19: A systematic review. *Diabetes Metab. Syndr. Clin. Res. Rev.* **14**, 655–659 (2020).
 182. Qin, R. *et al.* α 2,6-Sialylation is Upregulated in Severe COVID-19 Implicating the Complement Cascade. *medRxiv* 2022.06.06.22275981 (2022)
doi:10.1101/2022.06.06.22275981.
 183. Mellacheruvu, D. *et al.* The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nat. Methods* **10**, 730–736 (2013).
 184. Doherty, M., Buchy, P., Standaert, B., Giaquinto, C. & Prado-Cohrs, D. Vaccine impact: Benefits for human health. *Vaccine* **34**, 6707–6714 (2016).
 185. Iuliano, A. D. *et al.* Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* **391**, 1285–1300 (2018).
 186. Fry, A. M. *et al.* Modeling the effect of different vaccine effectiveness estimates on the number of vaccine-prevented influenza-associated hospitalizations in older adults. *Clin. Infect. Dis.* **59**, 406–409 (2014).
 187. World Health Organization. *Evaluation of influenza vaccine effectiveness: a guide to the design and interpretation of observational studies.* (World Health

- Organization, 2017).
188. Centers for Disease Control and Prevention. Estimated Influenza Illnesses, Medical visits, and Hospitalizations Averted by Vaccination in the United States — 2019–2020 Influenza Season. <https://www.cdc.gov/flu/about/burden-averted/2019-2020.htm> (2020).
 189. Potluri, T. *et al.* Age-associated changes in the impact of sex steroids on influenza vaccine responses in males and females. *npj Vaccines* **4**, Article number: 29 (2019) (2019).
 190. Tsang, J. S. *et al.* Improving Vaccine-Induced Immunity: Can Baseline Predict Outcome? *Trends Immunol.* **41**, 457–465 (2020).
 191. Tsang, J. S. *et al.* Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* **157**, 499–513 (2014).
 192. Poland, G. A., Ovsyannikova, I. G. & Jacobson, R. M. Immunogenetics of seasonal influenza vaccine response. *Vaccine* **26**, (2008).
 193. Ferrucci, L. & Fabbri, E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nat. Rev. Cardiol.* **15**, 505–522 (2018).
 194. Purdy, J. C. & Shatzel, J. J. The hematologic consequences of obesity. *Eur. J. Haematol.* **106**, 306–319 (2021).
 195. Pereira, B., Xu, X. N. & Akbar, A. N. Targeting Inflammation and Immunosenescence to Improve Vaccine Responses in the Elderly. *Front. Immunol.* **11**, 1–15 (2020).
 196. Diaz, A. *et al.* Metformin improves in vivo and in vitro B cell function in individuals

- with obesity and Type-2 Diabetes. *Vaccine* **35**, 2694–2700 (2017).
197. Selman, M. H. J. *et al.* Changes in antigen-specific IgG1 Fc N-glycosylation upon influenza and tetanus vaccination. *Mol. Cell. Proteomics* **11**, 1–10 (2012).
198. Wang, J. R. *et al.* Glycomic signatures on serum IgGs for prediction of postvaccination response. *Sci. Rep.* **5**, 1–11 (2015).
199. Wang, T. T. *et al.* Anti-HA Glycoforms Drive B Cell Affinity Selection and Determine Influenza Vaccine Efficacy. *Cell* **162**, 160–169 (2015).
200. Plomp, R. *et al.* Comparative glycomics of immunoglobulin A and G from saliva and plasma reveals biomarker potential. *Front. Immunol.* **9**, 1–12 (2018).
201. Roque-Barreira, M. C. & Campos-Neto, A. Jacalin: an IgA-binding lectin. *J. Immunol.* **134**, 1740 LP – 1743 (1985).
202. Audfray, A. *et al.* Fucose-binding lectin from opportunistic pathogen *Burkholderia ambifaria* binds to both plant and human oligosaccharidic epitopes. *J. Biol. Chem.* **287**, 4335–4347 (2012).
203. Wu, S. *et al.* Evaluation of determinants of the serological response to the quadrivalent split-inactivated influenza vaccine. *Mol. Syst. Biol.* **18**, 1–14 (2022).
204. Ryan, K. A. *et al.* Cellular immune response to human influenza viruses differs between H1N1 and H3N2 subtypes in the ferret lung. *PLoS One* **13**, e0202675 (2018).
205. Cao, Y. *et al.* Differential responses of innate immunity triggered by different subtypes of influenza A viruses in human and avian hosts. *BMC Med. Genomics* **10**, 41–54 (2017).

206. Belongia, E. A. *et al.* Repeated annual influenza vaccination and vaccine effectiveness: review of evidence. *Expert Rev. Vaccines* **16**, 723–736 (2017).
207. Martin, M. *et al.* Complement activation and plasma levels of C4b-binding protein in critical limb ischemia patients. *J. Vasc. Surg.* **50**, 100–106 (2009).
208. Giampaoli, O., Conta, G., Calvani, R. & Miccheli, A. Can the FUT2 Non-secretor Phenotype Associated With Gut Microbiota Increase the Children Susceptibility for Type 1 Diabetes? A Mini Review. *Front. Nutr.* **7**, 1–9 (2020).
209. Guo, M. *et al.* Distribution of Lewis and Secretor polymorphisms and corresponding CA19-9 antigen expression in a Chinese population. *FEBS Open Bio* **7**, 1660–1671 (2017).
210. Galeev, A. *et al.* The role of the blood group-related glycosyltransferases FUT2 and B4GALNT2 in susceptibility to infectious disease. *Int. J. Med. Microbiol.* **311**, 151487 (2021).
211. Rodríguez-Díaz, J. *et al.* Relevance of secretor status genotype and microbiota composition in susceptibility to rotavirus and norovirus infections in humans. *Sci. Rep.* **7**, 1–10 (2017).
212. Bucardo, F. *et al.* The Lewis A phenotype is a restriction factor for Rotateq and Rotarix vaccine-take in Nicaraguan children. *Sci. Rep.* **8**, 1–8 (2018).
213. Raza, M. W. *et al.* Association between secretor status and respiratory viral illness. *Br. Med. J.* **303**, 815–818 (1991).
214. Lindesmith, L. C. *et al.* Virus–Host Interactions Between Nonsecretors and Human Norovirus. *Cmgh* **10**, 245–267 (2020).

215. Lynn, D. J., Benson, S. C., Lynn, M. A. & Pulendran, B. Modulation of immune responses to vaccination by the microbiota: implications and potential mechanisms. *Nat. Rev. Immunol.* **0123456789**, (2021).
216. Philipp, R. *et al.* Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 (Secretor) genotype. *Proc. Natl. Acad. Sci.* **108**, 19030–19035 (2011).
217. Tong, M. *et al.* Reprogramming of gut microbiome energy metabolism by the FUT2 Crohn's disease risk polymorphism. *ISME J.* **8**, 2193–2206 (2014).
218. Borey, M. *et al.* Links between fecal microbiota and the response to vaccination against influenza A virus in pigs. *npj Vaccines* **6**, Article number: 92 (2021) (2021).
219. Xiao, L. *et al.* Human milk oligosaccharide 2'-Fucosyllactose improves innate and adaptive immunity in an influenza-specific murine vaccination model. *Front. Immunol.* **9**, 1–16 (2018).
220. Kim, Y.-J. *et al.* Complement C3 Plays a Key Role in Inducing Humoral and Cellular Immune Responses to Influenza Virus Strain-Specific Hemagglutinin-Based or Cross-Protective M2 Extracellular Domain-Based Vaccination. *J. Virol.* **92**, 1–14 (2018).
221. Ingels, H. *et al.* Immunodeficiency among Children with Recurrent Invasive Pneumococcal Disease. *Pediatr. Infect. Dis. J.* **34**, 644–651 (2015).
222. Gunput, S. T. G. *et al.* Complement activation by salivary agglutinin is secretor status dependent. *Biol. Chem.* **396**, 35–43 (2015).

223. Read, B. J. *et al.* Mannose-binding lectin and complement mediate follicular localization and enhanced immunogenicity of diverse protein nanoparticle immunogens. *Cell Rep.* **38**, 110217 (2022).
224. Tokatlian, T. *et al.* Innate immune recognition of glycans targets HIV nanoparticle immunogens to germinal centers. *Science (80-.).* **363**, 649–654 (2019).
225. Li, J., Hsu, H. C., Mountz, J. D. & Allen, J. G. Unmasking Fucosylation: from Cell Adhesion to Immune System Regulation and Diseases. *Cell Chem. Biol.* **25**, 499–512 (2018).
226. Nyström, K. *et al.* Virus-induced transcriptional activation of host FUT genes associated with neo-expression of Ley in cytomegalovirus-infected and sialyl-Lex in varicella-zoster virus-infected diploid human cells. *Glycobiology* **17**, 355–366 (2007).
227. Nyström, K. *et al.* Induction of sialyl-Lex expression by herpes simplex virus type 1 is dependent on viral immediate early RNA-activated transcription of host fucosyltransferase genes. *Glycobiology* **19**, 847–859 (2009).
228. Fan, R. *et al.* Alterations of fucosyltransferase genes and fucosylated glycans in gastric epithelial cells infected with helicobacter pylori. *Pathogens* **10**, 1–10 (2021).
229. Tinoco, R., Carrette, F., Henriquez, M. L., Fujita, Y. & Bradley, L. M. Fucosyltransferase Induction during Influenza Virus Infection Is Required for the Generation of Functional Memory CD4 + T Cells . *J. Immunol.* **200**, 2690–2702 (2018).

230. Memoli, M. J. *et al.* Evaluation of antihemagglutinin and antineuraminidase antibodies as correlates of protection in an influenza A/H1N1 virus healthy human challenge model. *MBio* **7**, e00417-16 (2016).
231. Ng, S. *et al.* Novel correlates of protection against pandemic H1N1 influenza A virus infection. *Nat. Med.* **25**, 962–967 (2019).
232. York, W. S. *et al.* GlyGen: Computational and Informatics Resources for Glycoscience. *Glycobiology* **30**, 72–73 (2020).
233. Carbon, S. *et al.* The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).
234. Mi, H. *et al.* PANTHER version 16: A revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* **49**, D394–D403 (2021).

Appendices

Appendix I. List of lectins printed on microarrays for serum glycomic analyses described in Chapter 2.

Probe	Source	Species of origin	Printing concentration. (mg/ml)	Inhibiting Sugar*
ConA	Vector Labs	<i>Concanavalin A</i>	2.0	Man
PHA-E	Vector Labs	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
AIA	Vector Labs	<i>Artocarpus integrifolia</i>	2.0	Gal
LcH	GlycoMatrix	<i>Lens culinaris</i>	2.0	Man
LcH	Vector Labs	<i>Lens culinaris</i>	2.0	Man
SNA	Vector Labs	<i>Sambucus nigra</i>	2.0	Lac
WGA	Vector Labs	<i>Triticum vulgare</i>	2.0	GlcNAc
VVA	Vector Labs	<i>Vicia villosa</i>	2.0	Gal
PNA	Vector Labs	<i>Arachis hypogaea</i>	2.0	Gal
AAL	Vector Labs	<i>Aleuria aurantia</i>	1.5	Fuc
PSA	Vector Labs	<i>Pisum sativum</i>	2.0	Man
PSA	GlycoMatrix	<i>Pisum sativum</i>	2.0	Man
UEA-I	Vector Labs	<i>Ulex europaeus</i>	2.0	Fuc
PHA-L	Vector Labs	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
LEL	Vector Labs	<i>Lycopersicon esculentum</i>	2.0	GlcNAc
GSL-I	Vector Labs	<i>Griffonia simplicifolia</i>	2.0	Gal

MPL	Vector Labs	<i>Machura pomifera</i>	2.0	Gal
BPL	Vector Labs	<i>Bauhinia purpurea</i>	2.0	Gal
MAA-I	Vector Labs	<i>Maachia amurensis</i>	2.0	Lac
STL	Vector Labs	<i>Solanum tuberosum</i>	2.0	GlcNAc
WFA	Vector Labs	<i>Wisteria floribunda</i>	2.0	Gal
DBA	Vector Labs	<i>Dolichos Biflorus</i>	2.0	Gal
SBA	Vector Labs	<i>Glycine max</i>	2.0	Gal
HHL	Vector Labs	<i>Hippeastrum hybrid</i>	2.0	Man
DSA	Vector Labs	<i>Datura stramonium</i>	2.0	Lac
ECA	Vector Labs	<i>Erythrina cristagalli</i>	2.0	GlcNAc
GSL-II	Vector Labs	<i>Griffonia simplicifolia</i>	2.0	GlcNAc
GSL-IB4	Vector Labs	<i>Griffonia simplicifolia</i>	2.0	Gal
LTL	Vector Labs	<i>Lotus tetragonolobus</i>	2.0	Fuc
GNA	Vector Labs	<i>Galanthus nivalis</i>	2.0	Man
NPL	Vector Labs	<i>Narcissus pseudonarcissus</i>	2.0	Man
MAA-II	Vector Labs	<i>Maachia amurensis</i>	1.0	Lac
UEA	Sigma	<i>Ulex europaeus</i>	2.0	Fuc
PHA-L	Sigma	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
MAA	Sigma	<i>Maachia amurensis</i>	2.0	Lac
PHA-E	Sigma	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
SNA	Sigma	<i>Sambucus nigra</i>	2.0	Lac
MNA-M	EY Labs	<i>Morniga M</i>	2.0	Man

PNA	EY Labs	<i>Sambucus nigra</i>	2.0	Gal
SNA-I	EY Labs	<i>Sambucus nigra</i>	2.0	Lac
SNA-II	EY Labs	<i>Sambucus nigra</i>	2.0	Gal
PHA-E	EY Labs	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
CA	EY Labs	<i>Colchicum autumnale</i>	2.0	GlcNAc
MNA-G	EY Labs	<i>Morniga G</i>	2.0	Gal
UEA-II	EY Labs	<i>Ulex europaeus</i>	2.0	GlcNAc
CSA	EY Labs	<i>Pure cytissus</i>	2.0	Gal
UDA	EY Labs	<i>Urtica dioica</i>	2.0	GlcNAc
LcH	EY Labs	<i>Lens culinaris</i>	2.0	Man
TL	EY Labs	<i>Tulipa sp.</i>	2.0	GlcNAc
ACA	EY Labs	<i>Amaranthus caudatus</i>	2.0	GlcNAc
Blackbean	EY Labs	<i>Phaseolus vulgaris</i>	2.0	Lac
PHA-L	EY Labs	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
UEA-I	EY Labs	<i>Ulex europaeus</i>	2.0	Fuc
VVA	EY Labs	<i>Vicia villosa</i>	2.0	Gal
LTL Lotus	EY Labs	<i>Lotus tetragonolobus</i>	2.0	Fuc
MAA	EY Labs	<i>Maachia amurensis</i>	2.0	Lac
PTA GalNAc	EY Labs	<i>Psophocarpus tetragonolobus</i>	2.0	Gal
AMA	EY Labs	<i>Arum maculatum</i>	2.0	Man
PTA Gal	EY Labs	<i>Psophocarpus tetragonolobus</i>	2.0	Gal

ASA	EY Labs	<i>Allium sativum</i>	2.0	Man
HPA	Sigma	<i>Helix pomatia</i>	2.0	Gal
AIA	GlycoMatrix	<i>Artocarpus integrifolia</i>	2.0	Gal
BPL	GlycoMatrix	<i>Bauhinia purpurea</i>	2.0	Gal
TJA-II	Aniara	<i>Trichosanthes japanica</i>	2.0	Lac
diCBM40	expressed in-house	<i>Clostridium perfringens</i>	1.5	Lac
AOL	TCI America	<i>Aspergillus oryzae</i>	1.5	Fuc
AOL	TCI America	<i>Aspergillus oryzae</i>	2.0	Fuc
Griffithsin	expressed in-house	<i>Griffithsia</i>	1.7	Man
SLBR-N	expressed in-house	<i>Streptococcus gordonii</i>	2.0	Lac
SLBR-H	expressed in-house	<i>Streptococcus gordonii</i>	2.0	Lac
SLBR-B	expressed in-house	<i>Streptococcus gordonii</i>	2.3	Lac
BanLec H84T	expressed in-house	<i>Musa paradisiaca</i>	1.0	Man
BamBL	expressed in-house	<i>Burkholderia cepacia</i>	1.5	Fuc
BC2L-A	expressed in-house	<i>Burkholderia cepacia</i>	0.6	Man
Protein A	Thermo Fisher		1.0	/
Protein G	Thermo Fisher		1.0	/
Protein L	Thermo Fisher		1.0	/
Anti-IgM	Thermo Fisher		/	/
Anti-IgM	Thermo Fisher		/	/

Anti-Lewis Y	Abcam		/	/
Anti-MUC5AC	Sigma		/	/
Anti-Lewis A	Abcam		/	/
Anti-Lewis B	Abcam		/	/
Anti-Forsman	Abcam		/	/
Anti-Lewis X	Sigma		/	/
Anti-Lewis B	Sigma		/	/
Anti-MBL	Abcam		/	/
Anti-ST6GAL1	Abcam		/	/
Anti-Galectin 9	R&D Systems		/	/
Anti-Galectin 3	Abcam		/	/
Anti-Galectin 1	Abcam		/	/
Anti-H2	SCBT		/	/
Anti-H1	Thermo Fisher		/	/
Anti-NEU1	Thermo Fisher		/	/
Anti-NEU3	MBL International		/	/
Anti-PolySia	Absolute Antibody		/	/
Cholera Toxin B	Sigma	<i>Vibrio cholerae</i>	2.0	Lac
RCA120	Vector Labs	<i>Ricinus communis</i>	2.0	Gal
Ricin B	Vector Labs	<i>Ricinus communis</i>	1.0	Lac

*Inhibiting sugars: Man, mannose; Gal, galactose; Fuc, fucose; GlcNAc, N-acetylglucosamine; Lac, lactose.

Appendix II. List of lectins printed on microarrays for plasma glycomic analyses described in Chapter 3.

Probe	Source	Species of Origin	Printing concentration (mg/ml)	Inhibiting Sugar*
ACA	EY	<i>Amaranthus caudatus</i>	2.0	GlcNAc
AIA	EY	<i>Artocarpus integrifolia</i>	2.0	Gal
AIA	Vector (lot#1)	<i>Artocarpus integrifolia</i>	2.0	Gal
AIA v2	Vector (lot#2)	<i>Artocarpus integrifolia</i>	2.0	Gal
AIA	GlycoMatrix	<i>Artocarpus integrifolia</i>	1.5	Fuc
Anti Lewis A	Abcam	/	as received	/
Anti Lewis B	Sigma	/	as received	/
Anti O-GlcNAc	Thermo Fisher	/	as received	/
Anti PolySia	Absolute Antibody	/	as received	/
Anti PSGL-1	Sigma	/	as received	/
Anti Sialyl Lewis X	GeneTex	/	as received	/
Anti VCAM-1	Sigma	/	as received	/
Anti VE cadherin	Thermo Fisher	/	as received	/
AOL	TCI America	<i>Aspergillus oryzae</i>	2.0	Fuc
BamBL	<i>Burkholderia cepacia</i>	expressed in-house	1.5	Fuc
BanLec H84T	UAlberta	<i>Musa paradisiaca</i>	1.0	Man
BPL	Vector	<i>Bauhinia purpurea</i>	2.0	Gal
ConA	Sigma	<i>Canavalia ensiformis</i>	2.0	Man
ConA	Vector	<i>Canavalia ensiformis</i>	2.0	Man
diCBM40	UAlberta	<i>Clostridium perfringens</i>	1.5	Lac
DSA	Vector	<i>Datura stramonium</i>	2.0	Lac
ECA	Vector	<i>Erythrina cristagalli</i>	2.0	GlcNAc

LcH	Aniara (lot #1)	<i>Lens culinaris</i>	1.5	Man
LcH	Aniara (lot #2)	<i>Lens culinaris</i>	1.5	Man
LcH	EY	<i>Lens culinaris</i>	2.0	Man
LEL	Vector	<i>Lycopersicon esculentum</i>	2.0	GlcNAc
MNA-M	EY	<i>Morniga M</i>	2.0	Man
MPL	Vector	<i>Machura pomifera</i>	2.0	Gal
PHA-E	EY	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
PHA-E	Vector	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
PHA-L	EY	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
Protein A	Thermo Fisher	/	0.5	/
Protein G	Thermo Fisher	/	1.0	/
PSA	Vector	<i>Pisum sativum</i>	2.0	Man
RCA120	Vector	<i>Ricinus communis</i>	2.0	Gal
SLBR-B	UAlberta	<i>Streptococcus gordonii</i>	2.3	Lac
SLBR-H	UAlberta	<i>Streptococcus gordonii</i>	2.0	Lac
SLBR-N	UAlberta	<i>Streptococcus gordonii</i>	2.0	Lac
SNA	Vector	<i>Sambucus nigra</i>	2.0	Lac
STL	Vector	<i>Solanum tuberosum</i>	2.0	GlcNAc
TJA-II	Aniara	<i>Trichosanthes japonica</i>	2.0	Lac
WGA	Vector	<i>Triticum vulgare</i>	2.0	GlcNAc

*Inhibiting sugars: Man, mannose; Gal, galactose; Fuc, fucose; GlcNAc, N-acetylglucosamine; Lac, lactose.

Appendix III. List of lectins printed on microarrays for ferret tissue glycomic analyses described in Chapter 3.

Probe	Source	Species of Origin	Printing concentration (mg/ml)	Inhibiting Sugar*
ConA	Millipore Sigma	<i>Concanavalin A</i>	2.0	Man
GNA	Millipore Sigma	<i>Galanthus nivalis</i>	2.0	Man
AIA	GlycoMatrix	<i>Artocarpus integrifolia</i>	1.5	Gal
PHA-E	Vector Labs	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
AIA	Vector Labs	<i>Artocarpus integrifolia</i>	2.0	Gal
LcH	Vector Labs	<i>Lens culinaris</i>	1.5	Man
SNA	Vector Labs	<i>Sambucus nigra</i>	2.0	Lac
WGA	Vector Labs	<i>Triticum vulgare</i>	2.0	GlcNAc
PNA	Vector Labs	<i>Arachis hypogaea</i>	2.0	Gal
AAL	Vector Labs	<i>Aleuria aurantia</i>	1.5	Fuc
PSA	Vector Labs	<i>Pisum sativum</i>	2.0	Man
VVA	Vector Labs	<i>Vicia villosa</i>	2.0	Gal
UEA-I	Vector Labs	<i>Ulex europaeus</i>	2.0	Fuc
LEL	Vector Labs	<i>Lycopersicon esculentum</i>	2.0	GlcNAc
MPL	Vector Labs	<i>Machura pomifera</i>	2.0	Gal
BPL	Vector Labs	<i>Bauhinia purpurea</i>	2.0	Gal

MAA-I	Vector Labs	<i>Maachia amurensis</i>	2.0	Lac
STL	Vector Labs	<i>Solanum tuberosum</i>	2.0	GlcNAc
WFA	Vector Labs	<i>Wisteria floribunda</i>	2.0	Gal
DBA	Vector Labs	<i>Dolichos Biflorus</i>	2.0	Gal
SBA	Vector Labs	<i>Glycine max</i>	2.0	Gal
HHL	Vector Labs	<i>Hippeastrum hybrid</i>	2.0	Man
DSA	Vector Labs	<i>Datura stramonium</i>	2.0	Lac
ECA	Vector Labs	<i>Erythrina cristagalli</i>	2.0	GlcNAc
GSL-II	Vector Labs	<i>Griffonia simplicifolia</i>	2.0	GlcNAc
LTL	Vector Labs	<i>Lotus tetragonolobus</i>	2.0	Fuc
GNA	Vector Labs	<i>Lotus tetragonolobus</i>	2.0	Man
NPL	Vector Labs	<i>Galanthus nivalis</i>	2.0	Man
MAA-II	Vector Labs	<i>Narcissus pseudonarcissus</i>	1.0	Lac
PSA	GlycoMatrix	<i>Maachia amurensis</i>	2.0	Man
UEA	Millipore Sigma	<i>Ulex europaeus</i>	2.0	Fuc
PHA-L	Millipore Sigma	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
MAA	Millipore Sigma	<i>Maachia amurensis</i>	2.0	Lac
SNA	Millipore Sigma	<i>Sambucus nigra</i>	2.0	Lac
MNA-M	EY Labs	<i>Morniga M</i>	2.0	Man
SNA-I	EY Labs	<i>Sambucus nigra</i>	2.0	Lac
SNA-II	EY Labs	<i>Sambucus nigra</i>	2.0	Gal
PHA-E	EY Labs	<i>Phaseolus vulgaris</i>	2.0	GlcNAc

CA	EY Labs	<i>Colchicum autumnale</i>	2.0	GlcNAc
MNA-G	EY Labs	<i>Morniga G</i>	2.0	Gal
UEA-II	EY Labs	<i>Ulex europaeus</i>	2.0	GlcNAc
CSA	EY Labs	<i>Pure cytisus</i>	2.0	Gal
UDA	EY Labs	<i>Urtica dioica</i>	2.0	GlcNAc
LcH	EY Labs	<i>Lens culinaris</i>	2.0	Man
TL	EY Labs	<i>Tulipa sp.</i>	2.0	GlcNAc
ACA	EY Labs	<i>Amaranthus caudatus</i>	2.0	GlcNAc
Blackbean	EY Labs	<i>Phaseolus vulgaris</i>	2.0	Lac
PHA-L	EY Labs	<i>Phaseolus vulgaris</i>	2.0	GlcNAc
UEA-I	EY Labs	<i>Ulex europaeus</i>	2.0	Fuc
LTL Lotus	EY Labs	<i>Lotus tetragonolobus</i>	2.0	Fuc
MAA	EY Labs	<i>Maachia amurensis</i>	2.0	Lac
PTA GalNAc	EY Labs	<i>Psophocarpus tetragonolobus</i>	2.0	Gal
AMA	EY Labs	<i>Arum maculatum</i>	2.0	Man
PTA Gal	EY Labs	<i>Psophocarpus tetragonolobus</i>	2.0	Gal
ASA	EY Labs	<i>Allium sativum</i>	2.0	Man
HPA	Millipore Sigma	<i>Helix pomatia</i>	2.0	Gal
TJA-II	Aniara	<i>Trichosanthes japanica</i>	2.0	Lac

diCBM40	UAlberta	<i>Clostridium perfringens</i>	1.5	Lac
AOL	TCI America	<i>Aspergillus oryzae</i>	2.0	Fuc
Griffithsin	UAlberta	<i>Griffithsia</i>	1.7	Man
SLBR-N	UAlberta	<i>Streptococcus gordonii</i>	2.0	Lac
SLBR-H	UAlberta	<i>Streptococcus gordonii</i>	2.0	Lac
SLBR-B	UAlberta	<i>Streptococcus gordonii</i>	2.3	Lac
AIA	EY Labs	<i>Artocarpus integrifolia</i>	2.0	Gal
BanLec H84T	UAlberta	<i>Musa paradisiaca</i>	1.0	Man
Cholera Toxin B	Millipore Sigma	<i>Vibrio cholerae</i>	2.0	Lac
RCA120	Vector Labs	<i>Ricinus communis</i>	2.0	Gal
Ricin B	Vector Labs	<i>Ricinus communis</i>	1.0	Lac
LcH	Aniara	<i>Lens culinaris</i>	2.0	Man
LcH	Aniara	<i>Lens culinaris</i>	1.0	Man
PA-III	Elicityl	<i>Pseudomonas aeruginosa</i>	2.0	Fuc
BC2L-A	Elicityl	<i>Burkholderia cepacia</i>	2.0	Man
Anti Lewis A	Abcam		/	/
Anti Lewis X	Millipore Sigma		/	/
Anti Lewis X	Millipore Sigma		/	/
Anti Lewis Y	Abcam		/	/
Anti Lewis Y	Abcam		/	/
Anti Lewis B	Millipore Sigma		/	/

Anti Sialyl Lewis X	GeneTex	/	/
Anti H1	Thermo Fisher	/	/
Anti H2	SCBT	/	/
Anti IgM	Abcam	/	/
Anti IgA	Abcam	/	/
Anti O-GlcNAc	Abcam	/	/
Anti O-GlcNAc	Thermo Fisher	/	/
Anti PolySia	Absolute Antibody	/	/
Anti Lewis A	Abcam	/	/
Anti Lewis B	Abcam	/	/
Anti B.G.A	Abcam	/	/
Anti B.G.A	Thermo Fisher	/	/
Anti B.G.B	Abcam	/	/
Anti B.G.B	Thermo Fisher	/	/
Protein A	Thermo Fisher	0.5	/
Protein G	Thermo Fisher	1.0	/
Protein L	Thermo Fisher	0.5	/

*Inhibiting sugars: Man, mannose; Gal, galactose; Fuc, fucose; GlcNAc, N-acetylglucosamine; Lac, lactose.

Appendix IV. List of lectins printed on microarrays for serum glycomic analyses described in Chapter 4.

Probe	Species of origin	Source	Printing concentration (mg/ml)	Inhibiting Sugar*
ConA	<i>Concanavalin A</i>	Millipore Sigma	2	Man
GNA	<i>Galanthus nivalis</i>	Millipore Sigma	2	Man
AIA	<i>Artocarpus integrifolia</i>	GlycoMatrix/Bio-world	2	Gal
PHA-E	<i>Phaseolus vulgaris</i>	Vector Labs	2	GlcNAc
AIA	<i>Artocarpus integrifolia</i>	Vector Labs	2	Gal
LcH	<i>Lens culinaris</i>	Vector Labs	2	Man
SNA	<i>Sambucus nigra</i>	Vector Labs	2	Lac
WGA	<i>Triticum vulgare</i>	Vector Labs	2	GlcNAc
PNA	<i>Arachis hypogaea</i>	Vector Labs	2	Gal
AAL	<i>Aleuria aurantia</i>	Vector Labs	1.5	Fuc
PSA	<i>Pisum sativum</i>	Vector Labs	2	Man
PSA	<i>Pisum sativum</i>	GlycoMatrix/Bio-world	2	Man
UEA-I	<i>Ulex europaeus</i>	Vector Labs	2	Fuc
LEL	<i>Lycopersicon esculentum</i>	Vector Labs	2	GlcNAc
MPL	<i>Machura pomifera</i>	Vector Labs	2	Gal
BPL	<i>Bauhinia purpurea</i>	Vector Labs	2	Gal
MAA-I	<i>Maachia amurensis</i>	Vector Labs	2	Lac
STL	<i>Solanum tuberosum</i>	Vector Labs	2	GlcNAc
WFA	<i>Wisteria floribunda</i>	Vector Labs	2	Gal
DBA	<i>Dolichos Biflorus</i>	Vector Labs	2	Gal
SBA	<i>Glycine max</i>	Vector Labs	2	Gal
HHL	<i>Hippeastrum hybrid</i>	Vector Labs	2	Man
DSA	<i>Datura stramonium</i>	Vector Labs	2	Lac

ECA	<i>Erythrina cristagalli</i>	Vector Labs	2	GlcNAc
GSL-II	<i>Griffonia simplicifolia</i>	Vector Labs	2	GlcNAc
LTL	<i>Lotus tetragonolobus</i>	Vector Labs	2	Fuc
GNA	<i>Galanthus nivalis</i>	Vector Labs	2	Man
NPL	<i>Narcissus pseudonarcissus</i>	Vector Labs	2	Man
MAA-II	<i>Maachia amurensis</i>	Vector Labs	1	Lac
VVA	<i>Vicia villosa</i>	Vector Labs	2	Gal
UEA	<i>Ulex europaeus</i>	Millipore Sigma	2	Fuc
PHA-L	<i>Phaseolus vulgaris</i>	Millipore Sigma	2	GlcNAc
MAA	<i>Maachia amurensis</i>	Millipore Sigma	2	Lac
SNA	<i>Sambucus nigra</i>	Millipore Sigma	2	Lac
MNA-M	<i>Morniga M</i>	EY Labs	2	Man
SNA-I	<i>Sambucus nigra</i>	EY Labs	2	Lac
SNA-II	<i>Sambucus nigra</i>	EY Labs	2	Gal
PHA-E	<i>Phaseolus vulgaris</i>	EY Labs	2	GlcNAc
CA	<i>Colchicum autumnale</i>	EY Labs	2	GlcNAc
MNA-G	<i>Morniga G</i>	EY Labs	2	Gal
UEA-II	<i>Ulex europaeus</i>	EY Labs	2	GlcNAc
CSA	<i>Pure cytisus</i>	EY Labs	2	Gal
UDA	<i>Urtica dioica</i>	EY Labs	2	GlcNAc
LcH	<i>Lens culinaris</i>	EY Labs	2	Man
TL	<i>Tulipa sp.</i>	EY Labs	2	GlcNAc
ACA	<i>Amaranthus caudatus</i>	EY Labs	2	GlcNAc
Blackbean	<i>Phaseolus vulgaris</i>	EY Labs	2	Lac
PHA-L	<i>Phaseolus vulgaris</i>	EY Labs	2	GlcNAc
UEA-I	<i>Ulex europaeus</i>	EY Labs	2	Fuc
LTL	<i>Lotus tetragonolobus</i>	EY Labs	2	Fuc
MAA	<i>Maachia amurensis</i>	EY Labs	2	Lac
PTA GalNAc	<i>Psophocarpus tetragonolobus</i>	EY Labs	2	Gal

AMA	<i>Arum maculatum</i>	EY Labs	2	Man
PTA Gal	<i>Psophocarpus tetragonolobus</i>	EY Labs	2	Gal
ASA	<i>Allium sativum</i>	EY Labs	2	Man
HPA	<i>Helix pomatia</i>	Millipore Sigma	2	Gal
TJA-II	<i>Trichosanthes japonica</i>	Aniara	2	Lac
diCBM40	<i>Clostridium perfringens</i>	expressed in-house	1.5	Lac
AOL	<i>Aspergillus oryzae</i>	TCI America	2	Fuc
Griffithsin	<i>Griffithsia</i>	expressed in-house	1.7	Man
SLBR-N	<i>Streptococcus gordonii</i>	expressed in-house	2	Lac
SLBR-H	<i>Streptococcus gordonii</i>	expressed in-house	2	Lac
SLBR-B	<i>Streptococcus gordonii</i>	expressed in-house	2.3	Lac
BamBL	<i>Burkholderia cepacia</i>	expressed in-house	1.5	Fuc
BanLec H84T	<i>Musa paradisiaca</i>	expressed in-house	1	Man
Protein A		Thermo Fisher	0.5	/
Protein G		Thermo Fisher	1	/
Protein L		Thermo Fisher	0.5	/
Anti Lewis A		Abcam	as received	/
Anti Lewis X		Millipore Sigma	as received	/
Anti Lewis Y		Abcam	as received	/
Anti Lewis B		Millipore Sigma	as received	/
Anti Sialyl Lewis X		GeneTex	as received	/
Anti H1		Thermo Fisher / Invitrogen	as received	/
Anti H2		Santa Cruz Biotechnology	as received	/
Anti PolySia		Absolute Antibody	as received	/
Anti Lewis A		Thermo Fisher / Invitrogen	as received	/
Anti Lewis B		Abcam	as received	/
Anti B.G.A		Abcam	as received	/
Anti B.G.A		Thermo Fisher / Invitrogen	as received	/
Anti B.G.B		Abcam	as received	/

Anti B.G.B		Thermo Fisher / Invitrogen	as received	/
Cholera Toxin B	<i>Vibrio cholerae</i>	Millipore Sigma	1	Lac
RCA120	<i>Ricinus communis</i>	Vector Labs	2	Gal
Ricin B	<i>Ricinus communis</i>	Vector Labs	1	Lac

*Inhibiting sugars: Man, mannose; Gal, galactose; Fuc, fucose; GlcNAc, N-acetylglucosamine; Lac, lactose.