UC Davis

UC Davis Electronic Theses and Dissertations

Title

Unraveling Glycosylation Alterations in Alzheimer's Disease through Multi-Omics Analysis

Permalink

https://escholarship.org/uc/item/6jv834sv

Author

Tang, Xinyu

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Unraveling Glycosylation Alterations in Alzheimer's Disease through Multi-Omics Analysis

By

XINYU TANG

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Nutritional Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Angela M. Zivkovic, Chair

J. Bruce German

Danielle G. Lemay

Committee in Charge

2024

Acknowledgement

In the pursuit of my doctoral degree, I have been incredibly fortunate to have been surrounded by people who have provided indispensable support and guidance to my academic and personal development. First, I would like to express my deepest gratitude to my parents for their unwavering emotional support and their commitment to financing my education. Their encouragement and belief in me have been the cornerstone of my academic journey.

Second, I am immensely grateful to my major professor, Dr. Angela Zivkovic, for her invaluable guidance and emotional support throughout my doctoral studies. Dr. Zivkovic's expertise in nutrition, critical thinking in scientific research, and steadfast encouragement have been essential in the successful completion of my research and the fostering of my self-assurance. She has not only served as an exemplary academic mentor but has also played a critical role as a trusted advisor in my life.

Furthermore, I extend my heartfelt appreciation to the members of my dissertation committee, Dr. Bruce German and Dr. Danielle Lemay, for their insightful advice, constructive feedback, and scholarly guidance to this dissertation, which have significantly enriched the quality of this work.

I also wish to acknowledge my labmates, Jack (Jingyuan) Zheng, Brian Hong, Joanne Agus, Dr. Jea-Woo Kang, and Dr. Trevor (Chenghao) Zhu, for their collaboration and assistance in my research. They are cordial and easy-going people who have created a relaxed and harmonious communication environment. Addressing specific biological questions with them has allowed me to refine my computational skills, and their expertise in nutrition, biology, and biotechnology has enabled a more profound interpretation of my results within the biological

context. I would also like to acknowledge our collaborators, Dr. Carlito Lebrilla and his lab members, including Jennyfer Tena, Ryan Schindler, and Armin Oloumi, for providing liquid chromatography-mass spectrometry (LC-MS/MS) based approaches to detect and quantify N-glycans, glycoproteins, and glycosphingolipids in brain homogenates, serum, and iPSC microglia. I would also like to express my gratitude to Dr. Lee-Way Jin, Dr. Izumi Maezawa, Dr. Jacopo Di Lucente, and Dr. Danielle Harvey for providing brain and iPSC microglia samples, assistance with q-PCR measurements, and guidance regarding experimental design and manuscript writing. The expertise and generosity of all collaborators have greatly enhanced the scope of this research.

Lastly, I would like to extend my gratitude to my dear friends, Zhongyue Yang and Yingxin Su, whose companionship and support have been a strength on this challenging academic journey. Special recognition is also extended to Trevor (Chenghao) Zhu, the first member of the Zivkovic Lab I encountered, who propelled me to the bioinformatics and computational biology field. He continuously provides me with valuable suggestions and assistance, both in the research and personal matters. To all who have contributed to this endeavor, thank you all for being part of this journey.

Abstract

Alzheimer's Disease (AD) is the most common cause of dementia, characterized by the extracellular deposition of β-amyloid plaques and intracellular neurofibrillary tangles resulting from hyperphosphorylated tau protein. AD has emerged as a looming concern in the U.S., with an estimated 6.7 million Americans projected to live with AD by 2023, and about 1 in 9 people aged 65 and older affected by Alzheimer's dementia. Unfortunately, there is still no effective treatment to reverse or stop AD progression. Glycosylation is a post-translational modification where glycans are attached to peptide chains or lipids. Emerging studies have demonstrated a close link between aberrant protein glycosylation and AD pathology. Given its complexity and diversity, glycosylation holds promise to understand the precise pathogenesis of AD from a novel perspective. Meanwhile, nutritional interventions hold potential to modify glycosylation patterns, as evidenced by the widespread use of oral monosaccharide administration in managing certain congenital disorders of glycosylation. Hence, it is pivotal to investigate the involvement of glycosylation in AD pathology and explore the possibility of nutritional management in modulating glycosylation patterns. This dissertation includes four chapters that highlight work focusing on glycosylation alterations in AD and dietary impacts on glycosylation.

The first chapter delved into profiling glycosylation alterations in the postmortem brains of AD individuals, focusing on alterations in glycosyltransferases and N-linked glycans. Leveraging publicly available RNA-seq datasets encompassing seven brain regions and incorporating 1724 samples, we identified differentially expressed glycosyltransferases, including MGAT1, B4GALT1, GALNTs, subsequently confirmed by qPCR in an independent set of human medial temporal cortex (MTC) samples. N-glycan-related changes predicted by differential expression of these glycosyltransferases were confirmed by liquid chromatography-

mass spectrometry (LC-MS)-based N-glycan analysis in the MTC (n=9 AD vs. 6 controls). Our findings indicated that MGAT1 and B4GALT1, which are involved in the formation of complex N-linked glycans and galactosylation, were upregulated in multiple brain regions. This upregulation was reflected in an increased abundance of corresponding N-glycans. Additionally, we identified transcription factors (TFs), including nutrient-stimulated TFs, regulating glycosyltransferases, indicating the potential of dietary components to impact glycosylation. Furthermore, heterogeneous differential expression of glycosylation isoenzymes underscored the necessity to investigate the cell-specific functions of these isoenzymes.

The second chapter focused on glycosylation alterations in microglia, innate immune cells in the central nervous system. AD pathology has highlighted the importance of microglia due to their role in orchestrating neuroinflammation and phagocytosis. We characterized glycosylation changes in human microglia following (Amyloid-β Oligomer) AβO and (lipopolysaccharide) LPS stimulation. Both treatments had distinct impacts on the N-glycan profile of human microglia, as evidenced by distinct clustering of AβO- and LPS-activated human microglia glycan signatures on the PCA plot. AβO treatment increased sialylated N-glycan abundance, coupled with the upregulation of genes encoding for sialyltransferases, ST3GAL2, ST3GAL4, and ST3GAL6 in AβO-activated microglia. In contrast, LPS stimulation increased fucosylation in N-glycans, alongside the upregulation of genes encoding for fucosyltransferases, including FUT4, and the downregulation of fucosidase, FUCA1. Additionally, LPS stimulation decreased the abundance of complex-type N-glycans, aligned with downregulation of mannosidase genes (MAN1A1, MAN2A2, MAN1C1). However, the glycosphingolipid (GSL) profile did not exhibit significant changes in either AβO- or LPS-activated microglia.

While protein glycosylation alterations observed in the brain and in microglia reveal potential mechanisms by which glycosylation is involved in AD, brain tissue samples are not useful for diagnosis. In the third chapter we explored the potential utility of serum glycoproteins as biomarkers for AD diagnosis. Utilizing serum samples from around 100 AD patients and 100 age- and gender-matched controls, we quantified glycopeptides of the most abundant serum proteins. The results indicate that among the differentially expressed glycopeptides, the majority belonged to proteins involved in immune function, including immunoglobulins and acute-phase proteins such as alpha-1-antitrypsin (A1AT), alpha-2-macroglobulin (A2MG), and complement C3 (CO3). Non-fucosylated and mono-fucosylated immunoglobulins G1 and G2 were key discriminators between control and AD individuals. Partial Least Square-Discriminant Analysis demonstrated the potential of these glycopeptides in distinguishing AD from normal controls, albeit with some overlap. Confounding factors, such as age, sex, BMI, and ethnicity, had impacts on glycopeptide abundances.

Given the pivotal role of glycosylation in AD, the final chapter investigated the impacts of dietary supplements on glycosylation patterns. We developed a targeted glycoproteomic method to monitor site-specific glycoprofiles and quantities of the most abundant HDL-associated proteins using Orbitrap LC-MS for (glyco)peptide target discovery and QqQ LC-MS for quantitative analysis. Using this glycopeptide quantification workflow, we conducted a pilot study to determine whether HDL protein glycoprofiles are altered in healthy human participants in response to dietary glycan supplementation. Twenty-two healthy adult men/women (age 18-45) with a Body Mass Index (BMI) range of 18.5-25 were randomized to four treatment groups: placebo (n=4), N-acetylglucosamine (n=6), Spirulina (n=6), and galactose (n=6). Blood samples were collected at baseline and after four weeks of supplementation. Peptides and glycopeptides

on HDL were quantified using the new workflow. The results demonstrated the potential of dietary intervention in modifying the HDL glycoprofile even within a short-term period. Further studies with larger sample sizes are required to validate diet-induced HDL glycoprofile alterations.

Table of Contents

ACKNOWLEDGEMENT ..... **II**

ABSTRACT ..... **IV**

CHAPTER 1: TRANSCRIPTOMIC AND GLYCOMIC ANALYSES HIGHLIGHT PATHWAY-SPECIFIC GLYCOSYLATION ALTERATIONS UNIQUE TO ALZHEIMER'S DISEASE **1**

ABSTRACT 2

INTRODUCTION 4

MATERIALS AND METHODS 6

RESULTS 9

DISCUSSION 16

CONCLUSION 24

LIST OF ABBREVIATIONS 25

DECLARATIONS 27

ACKNOWLEDGEMENTS 29

TABLES AND FIGURES 30

REFERENCES 62

CHAPTER 2: UNIQUE N-GLYCOSYLATION SIGNATURES IN AB OLIGOMER- AND LIPOPOLYSACCHARIDE-ACTIVATED HUMAN IPSC-DERIVED MICROGLIA **73**

ABSTRACT 74

INTRODUCTION 75

METHODS 77

RESULTS 85

DISCUSSION 90

DECLARATIONS 96

ACKNOWLEDGEMENTS 97

TABLES AND FIGURES 98

REFERENCES 109

CHAPTER 3: GLYCOSYLATION ALTERATIONS IN SERUM OF ALZHEIMER'S DISEASE PATIENTS SHOW WIDESPREAD CHANGES IN N- GLYCOSYLATION OF PROTEINS RELATED TO IMMUNE FUNCTION, INFLAMMATION, AND LIPOPROTEIN METABOLISM 118

ABSTRACT 119

INTRODUCTION 120

MATERIALS AND METHODS 122

RESULTS 126

DISCUSSION 131

DECLARATIONS 135

ACKNOWLEDGEMENTS 136

TABLES AND FIGURES 137

REFERENCES 152

CHAPTER 4: QUANTITATIVE GLYCOPROTEOMICS OF HIGH-DENSITY LIPOPROTEINS 156

ABSTRACT 157

INTRODUCTION 158

MATERIALS AND METHODS 160

RESULTS AND DISCUSSION 167

CONCLUSIONS 175

DECLARATIONS 176

ACKNOWLEDGEMENTS 177

TABLES AND FIGURES 178

REFERENCES 191

Chapter 1: Transcriptomic and glycomic analyses highlight pathway-specific glycosylation alterations unique to Alzheimer's disease

Xinyu Tang[1*], Jennyfer Tena[2*], Jacopo Di Lucente[3,4], Izumi Maezawa[3,4], Danielle J. Harvey[5], Lee-Way Jin[3,4], Carlito B. Lebrilla[2], Angela M. Zivkovic[1**]

[1]Department of Nutrition, University of California, Davis, Davis, CA, USA

[2]Department of Chemistry, University of California, Davis, Davis, CA, USA

[3]Department of Pathology and Laboratory Medicine, School of Medicine, University of California, Davis, Sacramento, CA, USA

[4]UC Davis MIND Institute, Sacramento, CA, USA

[5]Division of Biostatistics, Department of Public Health Sciences, School of Medicine, University of California, Davis, Davis, CA, USA

*These two authors have equal contribution

**Corresponding author: Angela M. Zivkovic, email: amzivkovic@ucdavis.edu, One Shields Ave, Department of Nutrition, University of California, Davis, Davis, CA, 95616, USA

Published manuscript with links to all supplemental files available at:

https://www.nature.com/articles/s41598-023-34787-4

ABSTRACT

Background: Glycosylation has been found to be altered in the brains of individuals with Alzheimer's Disease (AD). However, it is unknown which specific glycosylation-related pathways are altered in AD dementia.

Methods: Using publicly available RNA-seq datasets covering seven brain regions and including 1724 samples, we identified glycosylation-related genes ubiquitously changed in individuals with AD. Several differentially expressed glycosyltransferases found by RNA-seq were confirmed by qPCR in a different set of human medial temporal cortex (MTC) samples (n=20 AD vs. 20 controls). N-glycan-related changes predicted by expression changes in these glycosyltransferases were confirmed by mass spectrometry (MS)-based N-glycan analysis in the MTC (n=9 AD vs. 6 controls).

Results: About 80% of glycosylation-related genes were differentially expressed in at least one brain region of AD participants (adjusted p-values < 0.05). Upregulation of MGAT1 and B4GALT1 involved in complex N-linked glycan formation and galactosylation, respectively, were reflected by increased concentrations of corresponding N-glycans. Isozyme-specific changes were observed in expression of the polypeptide N-acetylgalactosaminyltransferase (GALNT) family and the alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase (ST6GALNAC) family of enzymes. Several glycolipid-specific genes (UGT8, PIGM) were upregulated. The critical transcription factors regulating the expression of N-glycosylation and elongation genes were predicted and found to include STAT1 and HSF5. The miRNA predicted to be involved in regulating N-glycosylation and elongation glycosyltransferases were hsa-miR-1-3p and hsa-miR-16-5p, respectively.

Conclusions: Our findings provide an overview of glycosylation pathways affected by AD and potential regulators of glycosyltransferase expression that deserve further validation and suggest that glycosylation changes occurring in the brains of AD dementia individuals are highly pathway-specific and unique to AD.

Keywords: Glycosyltransferase, Glycomics, RNA-seq, Brain, Alzheimer's Disease

INTRODUCTION

Alzheimer's Disease (AD) is the most common form of dementia in the elderly, characterized by the accumulation of extracellular amyloid-β (Aβ) and intracellular hyperphosphorylated tau, as well as the accumulation of cholesterol-containing lipid droplets [1,2]. However, treatments to reduce Aβ load have failed to stop AD progression [3,4], highlighting the need to discover new targets for preventing and treating AD.

The search for new therapeutic targets for AD has turned to glycobiology because of recent evidence showing major alterations in N- and O-glycosylation in the brain, serum, and cerebrospinal fluid (CSF) of AD patients [5–9]. Glycosylation affects biological processes across the major brain cell types. For example, in neurons, the number of glycosylation sites on neuronal and synaptic proteins is altered in AD brains, leading to aberrant neuronal adhesion and synaptic transmission [10]. Similarly, dysregulated N-glycosylation and sialylation patterns in astrocytes and microglia affect extracellular functions and inflammatory responses [10,11].

Over 200 glycosyltransferases are involved in glycan biosynthesis, including sulfotransferases. Mutations in these enzymes cause Congenital Disorders of Glycosylation (CDG), and over 80% of CDGs are associated with neurological abnormalities [12]. For example, defects in alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase (MGAT1) are associated with neurologic defects concurrent with caspase3 activation and neuronal apoptosis [13]. Moreover, the specificity and non-redundant functions of isoenzyme families are still poorly understood. For example, dysfunction in the polypeptide N-acetylgalactosaminyltransferase 2 (GALNT2) isoenzyme, which belongs to the GALNT family of glycosyltransferases that initiate mucin-type O-glycosylation, reduced Aβ production by reducing O-GalNAc formation of the amyloid precursor protein (APP) [14]. In contrast,

GALNT6 enzymatic activity caused reduced Aβ40 and Aβ42 generation with excess GalNAc-type O-glycosylation of APP [15]. Although several studies have explored the brain glycoproteome in AD[10,16], it is not clear which glycosylation pathways are specifically affected in AD. Furthermore, the expression profiles of glycosylation genes in different brain regions remain uncharacterized.

In this study, the aim was to conduct an exploratory analysis using publicly available RNA-seq data to provide an overview of glycosylation-related pathways altered in AD across multiple brain regions, and to determine whether some of the predicted alterations could be confirmed by qPCR and N-glycomic analysis in one of these brain regions. We performed differential expression analysis of glycosylation genes in AD participants compared to controls and participants with other neurological conditions using a harmonized data set including RNA-seq data from the Mayo, Mount Sinai Brain Bank (MSBB), and Religious Orders Study/Memory and Aging Project (ROSMAP) studies. The RNA-seq data include seven brain regions: the dorsolateral prefrontal cortex (DLPFC), frontal pole (FP), inferior frontal cortex (IFG), superior temporal cortex (STG), parahippocampal gyrus (PHG), temporal cortex (TCX), and cerebellum (CER). Furthermore, we explored whether a subset of the differentially expressed genes discovered in the RNA-seq analysis also showed altered expression by qPCR analysis in a separate set of samples from AD participants vs. controls. We explored whether some of the gene expression differences observed by RNA-seq and qPCR are also observed by mass spectrometry-based N-glycan analysis of a separate set of samples from AD participants vs. controls. Finally, potential regulators of glycosyltransferase gene expression were inferred using curated databases and computational tools.

MATERIALS AND METHODS

Public RNA-seq data

RNA-seq data were obtained from the RNA-seq Harmonization Study, which harmonized RNA-seq data from the MSBB, Mayo, and ROSMAP studies by re-processing the reads with a uniform pipeline (The sample information is summarized in **Table 1-S1-2** and the Supplementary Methods section in supporting information).

RNA-seq data normalization and differential expression analysis

Raw gene count data were filtered and normalized with the trimmed mean of M-values (TMM) method with the edgeR package [17] in R version 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). The filter step retains genes that have count-per-million (CPM) above 10 in 70% of samples. Only protein-coding genes were used for analysis. A negative binomial model with the quasi-likelihood (QL) F-test was applied to perform the differential expression analysis among diagnosis groups. The model for MSBB and ROSMAP was adjusted for sex, age, RNA Integrity Number (RIN), post-mortem interval (PMI), and batch. Since 34%-35% of PMI information and all batch information were unavailable in Mayo study, the model for Mayo did not include these two covariates. ANOVA-like test and pairwise comparisons were performed as needed. P-values were adjusted for multiple hypothesis testing using Benjamini & Hochberg (BH) with a threshold of adjusted p-value ≤ 0.05 (Supplementary Methods).

Glycosylation-related gene selection

A list of 219 glycosylation-related genes was obtained from a previous publication [18].

Human Brain Tissue Collection

Human postmortem brain samples were drawn from the biorepository of the UC Davis Alzheimer's Disease Research Center (ADRC). Written informed consent, including consent for autopsy, was obtained from study participants or, for those with substantial cognitive impairment, a caregiver, legal guardian or other proxy. Study protocols were reviewed and approved by the Institutional Review Board (IRB). For postmortem diagnosis, we followed the National Institute on Aging-Alzheimer's Association guideline for the neuropathologic assessment of AD [19]. The samples for qPCR were from the medial temporal cortex (MTC). The specimens for the N-glycomics were from the MTC, lateral prefrontal cortex (LPFC), and lateral cerebellar cortex (LCBC). The brain tissue used for the current study was snap frozen during autopsy and was stored at -80 °C before RNA extraction (see Supplementary Methods).

Quantitative Real-Time PCR

DEGs identified by RNA-sequencing were selected for confirmation by qPCR based on the following criteria: 1) differentially expressed in multiple brain regions, 2) high fold-change in one or more brain regions, and 3) involved in N-glycan, O-GalcNAc-glycan, glycolipid biosynthesis or elongation, branching, or capping process. Total RNA from brain tissue samples (n=20 AD vs. n=20 controls) was extracted using RNeasy® Plus Universal Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. The primer sequences used are listed in **Table 1-S3**. Gene expression was normalized to an endogenous reference gene, β-actin. Data were analyzed by the 2-ΔΔCt method. All experiments were performed in duplicate (see Supplementary Methods).

Cell Membrane Extraction for N-glycan Analysis

A total of 51 tissue samples from 20 participants (n=10 AD vs. n=10 controls) (**Table 1-S4**) were analyzed. Tissue samples were homogenized, resuspended and lysed. Lysates were

pelleted by centrifugation at 2000 x g for 10 min to remove the nuclear fraction and cell debris. The supernatant was centrifuged at 200,000 x g, 4˚C for 45 min in series to remove other nonmembrane subcellular fractions (see Supplementary Methods).

Release and Purification of N-Glycans

Proteins were suspended and denatured. Glycans were released using 2 μL (500,000 units/ml) of peptide N-glycosidase F (glycerol-free; New England BioLabs, cat. no. P0705L). Ultracentrifugation at 200,000 x g was used to separate the N-glycans and the membrane fraction. The released N-glycans were purified by solid-phase extraction. N-Glycans were eluted, dried and reconstituted in 30μl of water before mass spectrometric analysis (see Supplementary Methods).

Glycomic Analysis by LC-MS/MS

Purified brain N-glycans were analyzed using an Agilent nano-LC/chip Q-ToF MS system. Relative abundances were determined by integrating peak areas for observed glycan masses, averaging abundances from instrumental triplicates and normalizing to the summed peak areas of all glycans detected (see Supplementary Methods).

Transcription Factors and miRNA-target Interaction Inference

Transcription factors (TF) and their binding motifs of glycosyltransferases were identified using the RcisTarget R package [20]. We used the motif collection based on human hg38 genome assembly and RefSeq genes and searched 10kbp around the transcription start site (TSS). MicroRNAs and their interactions with glycosyltransferases were inferred based on three databases with experimentally validated miRNA-target interactions: miRecords, miRTarBase,

and TarBase. The information was retrieved using the multiMiR package [21]. (see Supplementary Methods).

Statistical Analysis

For normally distributed data, groups were compared using Student's t-test. The fractions of subtypes, sialylated and fucosylated N-glycans were calculated as the sum of relative abundances of single glycans. Since we obtained more than one brain tissue from the same participant, differential abundance analysis was performed as linear mixed models using the lmerTest package in R[22] as shown here: glycan abundance ~diagnosis*region+(1|subject_id), where subject-specific effects were modeled as random effects. Contrasts between AD and control diagnosis were evaluated for each brain region. These contrasts produced t-statistics and corresponding p-values.

RESULTS

Glycosyltransferase profiling of human AD brain specimens

Human postmortem brain RNA-seq data covered seven distinct regions: the DLPFC, FP, IFG, STG, PHG, TCX, and CER. To identify the AD-associated transcriptome, we performed differential expression analysis on all protein-coding genes in each region independently, comparing AD participants to controls. Genes that did not meet the filtering criteria were excluded from differential expression analysis and included 13-17 genes in each brain region, representing approximately 5% of glycosylation-related genes in each region (**Table 1-S5**). The number of DEGs in each brain region varied with approximately 13822 protein-coding genes differentially expressed in at least one brain region (adjusted p-value < 0.05, **Figure 1-S1A**). Glycosylation-related genes represented 0.9%-1.3% of total DEGs, and about 80% of

glycosylation-related genes were differentially expressed in at least one brain region. The DLPFC and FP had fewer differentially expressed glycosyltransferases than other brain regions (**Figure 1-S1B**). Two upregulated genes, including alpha-N-acetylgalactosaminide alpha-2,6-saialyltransferase 2 (ST6GALNAC2), and three downregulated genes, including ST6GALNAC5 were observed in all brain regions (**Figure 1-S1C**). Genes encoding enzymes for O-glycosylation and sialylation had stronger signals than N-glycosylation and fucosylation in multiple brain regions of AD participants compared to controls (**Figure 1-S1D**).

Gene expression of MGAT1 was increased, with corresponding increases in complex N-glycans

We first analyzed the expression of genes involved in N-glycosylation. We detected 27 genes encoding glycosyltransferases for N-glycan synthesis in at least one brain region (**Figure 1-S2A**). Among these glycosyltransferases, MGAT1, essential for converting high-mannose structures to hybrid and complex glycans, had increased expression in the brains of AD participants compared to controls in the STG, TCX and CER, with the largest increase across studies in the TCX (**Figure 1-1A, Figure 1-S2A**). The increased gene expression of MGAT1 determined by RNA-seq was also observed by qPCR in an independent set of samples, showing approximately a 4-fold increase in gene expression in AD participants vs. healthy controls (**Figure 1-1B**).

We further sought to determine whether the increase in MGAT1 gene expression could also be observed as changes in the N-glycome in an additional independent set of brain samples. The N-glycomic analysis identified 429 N-glycans in the LPFC, MTC, and LCBC, with complex N-glycans being the most abundant structures in the human brain samples. The results showed that the percent abundance of complex N-glycans significantly increased in AD participants compared to controls in the MTC, and high-mannose N-glycans significantly decreased. The

observed increases in the final product of the pathway (i.e. complex N-glycans), and decreases in the initial precursor of the pathway (i.e. high-mannose N-glycans) are consistent with an increased flux through this pathway, of which the expression of the MGAT1 enzyme was observed to be increased by RNA-seq and qPCR (**Figure 1-1C**). The percent abundances of two intermediates of the pathway, the direct substrate for the MGAT1 enzyme, Hex:5 HexNAc:2 (Intermediate 1 in **Figure 1-1D-E**), and the direct product of the MGAT1 enzyme, Hex:5 HexNAc:3 (Intermediate 2 in **Figure 1-1D-E**), were both decreased in AD participants compared with controls. These data suggest a more complex regulation of flux through the entire pathway than a simple decrease in the immediate precursor and increase in the immediate product that would have been predicted by the increased expression of MGAT1 alone. Together these findings support an upregulation of complex N-glycan synthesis in the MTC of AD participants. On the other hand, although we found increased MGAT1 expression in the CER, no glycan subtype nor intermediate glycan was significantly changed in the LCBC in AD vs. controls (**Figure 1-S2B-D**).

Increased expression of B4GALT1 was accompanied by increased galactosylation of N-glycans in AD brains

The glycosylation process also involves elongation and branching biosynthetic steps. However, most glycosyltransferases for these steps belong to isoenzyme families with poorly understood non-redundant functions. We detected 17 glycosyltransferases for elongation and branching (**Figure 1-S3A**). β-1,4-galactosyltransferase 1 (B4GALT1) adds terminal galactose to N-glycans to generate type 2 LacNAc units (**Figure 1-2D**). The gene expression of B4GALT1 was substantially increased in AD participant brains in TCX (**Figure 1-2A, Figure 1-S3A**) and was not significantly changed in PA and PSP individuals, indicating that B4GALT1 may be an

AD-specific target gene (**Figure 1-2A**). The increased expression of B4GALT1 was also observed by qPCR in an independent set of tissue samples (**Figure 1-2B**), showing approximately a 2-fold increase in gene expression in AD participants compared with controls. According to the glycomic data, the percent abundance of non-galactosylated N-glycans significantly decreased in AD participants in the MTC (**Figure 1-2C**). Meanwhile, the proportion of N-glycans with six galactoses significantly increased in the MTC (**Figure 1-S3D**), consistent with an increase in flux through the galactosylation pathway predicted by the increase in B4GALT1 expression detected by RNA-seq and qPCR. The terminal galactose residues on N-glycans can be further sialylated. The N-glycomic results also demonstrated increased total sialylation in the MTC from AD participants (**Figure 1-S3E**). The increased sialylated N-glycans were mainly mono-sialylated. No significant N-glycan sialylation changes were observed in the LPFC and LCBC.

In addition to B4GALT1, the other two galactosyltransferases, B3GALT2 and B3GALT5, also significantly changed in AD in multiple brain regions, with results repeated by qPCR in the MTC (**Figure 1-S3B-C**). β-1,3-GalTs generate type-1 LacNAc units, which are relatively high in O-glycans and glycosphingolipids of humans.

Heterogeneous differential expression of isoenzymes involved in O-GalNAc-glycan initiation and alpha-2,6-sialylation on GalNAc residues

We observed heterogeneous differential expression within several notable isoenzyme families. For example, in the GALNT family of enzymes that initiate O-GalNAc-glycan biosynthesis (**Figure 1-3E**), the gene expression of GALNT15 and GALNT10 increased, but the expression of GALNT11 and GALNT17 decreased in the cortical regions but not in the CER (**Figure 1-3A, C, Figure 1-S4A-C**) according to the RNA-seq analysis. Notably, the increased

expression of these GALNT genes was unique to AD except for GALNT11 as they were not significantly different or changed in the opposite direction in MCI, PA, and PSP participants (**Figure 1-3A, C, Figure 1-S4A-C**). qPCR-based expression analysis in an independent sample set of tissues showed a 2-fold increase in gene expression of GALNT15 and GALNT10 but not GALNT11 in AD participants compared with controls (**Figure 1-3B, D, Figure 1-S4D**).

According to the RNA-seq results, in the ST6GALNAC family of enzymes, which transfer sialic acid onto N-acetylgalactosamine (GalNAc) in an alpha-2,6 linkage (**Figure 1-4G**), ST6GALNAC2 and ST6GALNAC3 expression increased in AD participants, while ST6GALNAC5 decreased (**Figure 1-4A, C, E**). ST6GALNAC3 and ST6GALNAC5 gene expression alterations were unique to AD participants and were also observed by qPCR in an independent set of samples (**Figure 1-4A-F**). Notably, the expression level of ST6GALNAC5 was much lower in the CER than in the six cortical regions across all diagnosis groups (**Figure 1-4E**). The gene expression of other isoenzyme families, such as the alpha-2,3-sialyltransferases (ST3GALs), alpha-2,6-sialyltransferases (i.e., ST6GALs), alpha-2,8-sialyltransferases (ST8SIAs), and fucosyltransferases (FUTs), were also altered in AD participants in multiple brain regions, with both increases and decreases in gene expression within the families. However, these results were either not confirmed by qPCR or observed in fewer brain regions (**Figure 1-S5A-D**).

Increased gene expression of key enzymes for GPI-anchor and galactocerebroside biosynthesis in AD participant brains

In addition to glycosylation on proteins, we analyzed the enzymes specific to lipid glycosylation. Accroding to the RNA-seq results, the temporal regions had more altered genes than the frontal regions (**Figure 1-S6A**). Among 14 detected lipid glycosylation genes,

phosphatidylinositol glycan anchor biosynthesis class M (PIGM) and UDP glycosyltransferase 8 (UGT8) increased in DLPFC and PHG of AD participants compared with controls (**Figure 1-5A, D, Figure 1-S6A**). qPCR results confirmed a 2-fold increase in the expression of these genes in AD participants (**Figure 1-5B, E**). Other critical genes were also upregulated in AD, such as PIGV (**Figure 1-S6A**). Since PIGM and PIGV transfer the first and second mannose to the GPI backbone, increased expression may indicate increased biosynthesis of GPI-anchored structures in the brains of AD participants (**Figure 1-5C**). Of the glycosphingolipid biosynthesis pathways, since UGT8 catalyzes a crucial step of galactocerebroside biosynthesis by transferring galactose to ceramide (**Figure 1-5F**), increased UGT8 expression may cause an increase in galactocerebroside structures in cortical brain regions in AD participants.

Several genes involved in glucosylceramide biosynthesis were also changed in AD participants across multiple brain regions and confirmed by qPCR, including alpha 1,4-galactosyltransferase (A4GALT) and B4GALT6 (p-value=0.05) (**Figure 1-S6B, D**).

Transcriptional and post-transcriptional regulation of glycosyltransferase genes

TFs and miRNAs regulate genes at the transcriptional and post-transcriptional level. We identified TF binding motifs over-represented on glycosyltransferase gene sets. The TF STAT1 had the highest Normalized Enrichment Score (NES) (NES = 5.28) for N-glycan core structure biosynthesis. MGAT1, MGAT3, MGAT4B, FUT8 and other genes were highly enriched for this TF binding motif (**Table 1-S6, Figure 1-S7A**). ZNF595, E2F7, ELK3, FOXO1, RXRA and other TFs were also identified as potential transcriptional regulators interacting with genes in the N-glycosylation pathway. The TF HSF5 was the most significant regulator (NES = 6.01) for the elongation and branching pathways. Eight genes including B3GALT2, B3GALT5 were highly enriched for this binding motif (**Table 1-S6, Figure 1-S7B**).

Using databases containing experimentally validated miRNA-target interactions, we identified miRNAs that possibly regulate various glycosylation pathways (**Table 1-S7**). For N-glycan biosynthesis, hsa-miR-1-3p and hsa-miR-124-3p regulated 21 and 16 out of 28 glycosyltransferases in this pathway, respectively. The targets of both miRNAs included MGAT1, MGAT2, MGAT4A/B, FUT8, and other genes (**Table 1-S7, Figure 1-S8A**). For elongation and branching, hsa-miR-16-5p (n=7) and hsa-miR-34a-5p (n=6) regulated the highest number of glycosyltransferases in the pathway. B4GALT1 and other genes were regulated by these miRNAs (**Table 1-S7, Figure 1-S8B**).

DISCUSSION

The brain is a complex organ with multiple brain regions performing distinct functions, such as memory, motor function, language, and learning, making it difficult to disentangle the pathogenesis and pathology of AD. Integrating publicly available RNA-seq data with qPCR and LC-MS/MS-based glycomics in separate sample sets, we performed a large-scale overview of the major alterations in N- and O-glycosylation, as well as lipid glycosylation in the context of AD across multiple human brain regions. We further identified candidate TFs and miRNAs that regulate the expression of glycosyltransferases for different glycosylation pathways.

In this study, we identified 13,882 protein-coding genes differentially expressed in at least one brain region, with the highest number of genes observed in TCX (n = 9271) and CER (n = 8113), followed by PHG (n = 4545) and IFG (n = 3275), and the lowest number of DEGs in FP (n = 154). The DEGs identified in each brain region were found to be unique, especially in TCX and CER, which had 2096 and 2396 unique DEGs, respectively (**Figure 1-S1A**). Our results are comparable with previous studies that reported the presence of 2174, 559, and 71 DEGs in the prefrontal cortex (PFC), precuneus (PRE), and visual cortex (VIC), respectively [23,24]. These cortical regions range from severely to mildly affected by tau pathology and neuronal loss. We found that 2027 DEGs in PFC overlapped with one or multiple brain regions in our study (**Figure 1-S9**). A previous analysis using transcriptomic data from the ROSMAP, MSBB, and Mayo data sets revealed 4485 DEGs with a fixed meta-analysis model [25].

Among the DEGs, in this study we specifically focused on glycosyltransferase genes. Systematic studies on the differential expression of glycosylation-related genes in AD are limited. However, several papers focused on the glycosyltransferases involved in specific pathways. A previous study reported an increase in GALNT4/6/7/8/10 expression in AD [15],

while another study found upregulation of STT3B, MAGT1, ALG2/6/12, and downregulation of RPN2, DAD1, OST4, OSTC, TUSC3, ALG3/9/12 in AD [10].

In this study, an increase in flux through the complex N-glycan synthesis pathway which could be predicted by an increase in the expression of the MGAT1 enzyme by both RNA-seq and qPCR, was corroborated by N-glycomic results showing decreases in high-mannose structures and concomitant increases in complex N-glycans in an independent set of brain tissues, procured from a separate cohort of AD participants vs. controls. High-mannose N-glycans are usually considered the precursor of hybrid and complex N-glycans. In the Golgi apparatus, the high-mannose structure is further trimmed and modified to form hybrid and complex N-glycans. MGAT1 initiates the biosynthesis of hybrid and complex structures. Decreased high-mannose N-glycans were also observed in the cerebrospinal fluid (CSF) of both males and females with AD [26]. Some studies have reported that complex N-glycans are essential for normal neurological development as MGAT1 defective mouse embryos failed in neural tube formation or died within eight weeks of birth [13,27]. Neuron-specific loss of MGAT1 function resulted in reduced complex and increased high-mannose N-glycans in hippocampal regions and in the spinal cord, concurrent with neuronal apoptosis and astrogliosis [13]. The most abundant N-glycan subtype observed in human brains in this study was complex N-glycans (~75% of total measured structures), which confirmed some previous studies [28]. However, other studies using different glycomic analytical approaches have reported the most abundant structures found were high mannose structures [7,29].

Our findings also demonstrated increased galactosylated N-glycans in AD participant brain samples, which would be consistent with an increase in flux through the galactosylation pathway predicted by both RNA-seq and qPCR. B4GALT1 is a galactosyltransferase enzyme

which adds galactose residues to N-glycans. Although few studies have investigated the role of B4GALT1 in AD, in rats after spinal cord injury, B4GALT1 protein partly colocalized with microglia, oligodendrocytes, and astrocytes and mediated pro-inflammatory reactions [30]. Conversely, B4GALT1 knockdown in mice resulted in an inhibition of glioblastoma development and increased survival [31]. Together this evidence suggests that galactosylation, particularly galactosylation mediated by B4GALT1, may be involved in regulating pro- vs. anti-inflammatory events in the central nervous system.

Studies using knockout models of glycogenes in animals and cell lines have demonstrated the presence of specific glycogenes can cause changes in the glycome [32,33]. Although direct evidence for specific regulation of the glycome by the expression of a single glycosyltransferase is limited, transcriptional regulation of the enzymes directly involved in glycan biosynthesis is likely to be the easiest way to study the regulation of the glycome [34,35]. In a previous glycomic study, the authors identified 62 glycosylation genes differentially expressed between the cortex and cerebellum and reported the correlation between these genes and glycomic patterns[29]. For example, the expression of Mgat3 was higher than the expression of other Mgat enzymes, consistent with a high abundance of bisecting N-glycans and a low abundance of complex, branched N-glycans. Moreover, Zhang, Q. et al. conducted a glycoproteomic study in which they compared their N-glycoproteomic data to four different transcriptome datasets for brain tissues [10]. They suggested that altered expression of enzymes in dolichol-oligosaccharide biosynthesis could contribute to the observed protein N-glycosylation changes in the AD brain.

An additional finding in this study is that isoenzyme-specific changes in enzymes that initiate O-glycosylation and sialylation, particularly those involving the GALNT and ST6GALNAC isoenzyme families respectively, may be involved in AD. Although the

isoenzymes belonging to these isoenzyme families catalyze the same reactions (i.e. all the GALNTs initiate O-glycosylation and all the ST6GALNACs add an alpha-2,6 sialic acid), several studies have reported their specificity for protein substrates in SimpleCell lines [36–38], emphasizing the need to study the non-redundant functions of these enzymes in different cell types under different conditions. The limited knowledge of isoenzyme non-redundant function also partly hampers the ability to correlate the abundance changes of glycan structures with differential expression. On the other hand, although most GALNTs and some ST6GALNACs have been studied in terms of their structure and enzyme kinetics [39–43], their physiological role in the central nervous system is still limited. Furthermore, their protein targets have not been defined. The heterogeneous AD-associated differential expression of these isoenzymes suggests a precisely regulated, pathway-specific process, with specific isoenzymes acting on different cell types, proteins, or glycosylation sites.

In addition to ST6GALNACs for alpha-2,6 sialylation, the capping process also involves alpha-2,3 sialyltransferases, alpha-2,8 sialyltransferases, and fucosyltransferases. In this study, the gene expression of these other sialylatransferases and fucosyltransferases was not as highly differential or was not changed in as many brain regions as the ST6GALNACs in AD participant brains. Sialylation has been extensively investigated in neuroinflammation. CD33 is a sialic acid-binding immunoglobulin-like lectin (Siglec) expressed in microglia in the brain [44]. The CD33 gene is a genetic risk factor for AD in GWAS studies, with a rare allele leading to decreased expression of CD33M (major isoform) and increased expression of CD33m (minor isoform), which is associated with protection from AD [45,46]. CD33 overexpression in the human brain is associated with impaired cognition [45,47,48]. When alpha-2,3 and, preferentially, alpha-2,6 sialylated glycans bind CD33, the Aβ phagocytosis activity of microglia increases [11,49,50].

Importantly, ST6GALNACs only modify GalNAc residues found on O-glycans and glycosphingolipids, highlighting the need for future studies focused on the sialylation changes in O-glycans and glycosphingolipids in AD.

Additionally, we observed alterations in glycosphingolipid pathways, where gene expression of UGT8, PIGM, and PIGV for galactocerebroside and GPI-anchor biosynthesis changed in DLPFC and PHG. In Parkinson's disease glycosphingolipid could control neuroinflammation and its metabolism was deregulated with decreased glucocerebrosidase activities in the brain of participants compared to control [51,52]. UGT8 initiates the biosynthesis of galactosylceramides. However, most glycosphingolipids in the mammalian brain derive from glucosylceramides, the precursor of complex gangliosides, including the four major gangliosides in the brain, GM1, GD1a, GD1b, and GT1b [52,53]. The changes in glycosphingolipid patterns are cell-specific as glucosylceramide-derived gangliosides are enriched in neurons, and galactosylceramides are enriched in oligodendrocytes and myelin [52–54]. The increased gene expression of UGT8 observed in this study highlights the involvement of oligodendrocytes and myelin in AD pathogenesis. In addition, in this study we identified increased gene expression of both PIGA and PIGM, which are both involved in GPI anchor formation, in the brains of AD participants. The GPI anchor attaches the immunoglobulin superfamily of cell adhesion modules (IgSF CAMs) to plasma membranes on neurons and functions in synapse formation, synaptic plasticity, and cell-to-cell communication[55,56]. Our findings support further research to better understand the involvement of glycolipid pathways in the pathogenesis of AD.

Finally, we identified several candidate regulators, including TFs and miRNAs that potentially regulate glycosyltransferase gene expression. STAT1 was a critical TF for

glycosyltransferases, regulating MGAT1, MGAT3, FUT8, and other genes in the N-glycosylation pathway. STAT1 is one of the key regulators for inflammatory activation within microglia in AD brains [57,58]. Its expression impaired spatial learning and memory in rats and could be induced by Aβ in a dose-dependent manner [59]. Our study also identified hsa-miR-1-3p, hsa-miR-124-3, hsa-miR-16-5p, and hsa-miR-34a-5p as the top candidate miRNA regulators for genes involved in N-glycosylation and elongation. hsa-miR-124 and hsa-miR-34a are two of the most abundantly expressed miRNAs in mammalian brains. Studies have demonstrated that hsa-miR-124-3p suppressed B4GALT1 protein translation in multiple cell lines [60]. Moreover, the expression of hsa-miR-124 gradually decreased in AD , which led to up-regulated BACE1 expression, ultimately responsible for Aβ production [61,62]. hsa-miR-34a has been reported to regulate ALG13, FUT8, GALNT7, and ST3GAL5 in neuroblastoma and hepatocellular carcinoma [63–66]. It ubiquitously participates in neurodevelopment and neuropathological processes and is dysregulated in AD. Its overexpression in animal models of AD was associated with neuronal death, and the suppression of hsa-miR-34a was neuroprotective [67]. In addition, hsa-miR-34a-5p has been found to be dysregulated in AD and other neurodegenerative diseases [68]. hsa-miR-16-5p was upregulated by Aβ in AD and induced neuronal apoptosis [69]. Our analysis was able to capture glycosyltransferase genes and their putative transcriptional and posttranscriptional regulators. These findings highlight the key glycosylation-related pathways and their regulators that could be important for the discovery of potential new targets for the management, treatment and prevention of AD. However, these findings need to be further validated in future experiments.

In summary, we have developed a system-level view of alterations in glycosylation-related genes in multiple brain regions in the context of AD and identified putative regulators of

glycosyltransferases from different pathways. Strengths of the study include investigation of RNA-seq data from multiple well-characterized AD cohorts spanning multiple brain regions, and the fact that several of the differential gene expression results found by RNA-seq are supported by qPCR in an independent set of brain samples. Given that typically 15-20% of protein-coding genes are non-concordant between RNA-seq and qPCR even within the same samples, our orthogonal findings in an independent sample set strengthen the likelihood that the observed changes may be important in AD [70]. We further studied the N-glycome with MS-based analysis in another independent set of samples. As expected, the transcriptomic alterations do not completely agree with the glycosylation changes observed by LC-MS/MS-based glycan analysis. This discrepancy may be caused by several other factors that potentially contribute to glycomic changes, including alterations in the expression of the most abundant glycoproteins in the brain, post-transcriptional and post-translational modifications of glycosyltransferase genes, donor sugar availability, competition among enzymes, co-factors, pH, chaperones and glycosidases, as well as glycome measurement methods. However, the fact that we found strong support for the transcription-level changes by measuring glycans directly in brain tissue homogenates indicates that these changes in glycosylation are consistent and important in AD etiology. There are also several critical limitations of this study. First, and most importantly, cellular heterogeneity in gene expression can be masked with the bulk RNA-seq approach. It is well appreciated that brain tissue homogenates from different development periods and regions of the brain comprise different relative abundances of cell types [71]. Furthermore, because neurons and oligodendrocytes represent the majority of cells in any given brain tissue homogenate, bulk RNA-seq, qPCR and glycomic analysis is most likely to yield information about these most abundant cells, and changes in less abundant cell types (e.g. microglia) are most likely to be

masked all together. Even further changes can be apparent between different subsets or phenotypes of cells (e.g. glutaminergic vs. GABAergic neurons, disease-associated microglia vs. resting microglia), and these also cannot be observed with the bulk approach in this study. Future studies are needed to address these important differences, including single-cell RNA-seq analysis, and various approaches to enrich specific cell types prior to gene expression and/or glycan analysis. Additionally, the differences in glycan profiles using different glycomic methods should be considered [72]. Future glycoproteomic studies are needed to investigate the abundance changes of glycosylated proteins to confirm whether the glycan changed observed here are due to the changes of the most abundant glycoproteins carrying those glycans. Site-specific glycoprofiles are also required in addition to the overall glycosylation patterns of cleaved glycans, since the glycosylation of proteins at different sites affect different aspects of protein function, metabolism, and structure. The association between sialylation, in particular alpha-2,6-sialylation, and neuroinflammation mediated by siglecs expressed in microglia and other cell types may also be an important future area of research. The findings of our current study, which provides a broad overview of key alterations in glycosylation pathways in the brain in AD participants and their putative regulators, support further research of glycosylation-related pathways to understand the pathology of AD.

CONCLUSION

The present study implies pathway-specific glycosylation alterations in the brains of AD dementia individuals. Differentially expressed glycosyltransferases mediate the aberrant brain glycan profile. And the putative transcriptional and post-transcriptional regulators for glycosyltransferases involved in the N-glycan core structure biosynthesis, elongation, and other pathways potentially facilitate the future development of AD therapeutics.

LIST OF ABBREVIATIONS

AD: Alzheimer's Disease

MCI: Mild Cognitive Impairment

NCI: No Cognitive Impairment

MGAT1: alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase

B4GALT: β-1,4-galactosyltransferase 1

GALNT: polypeptide N-acetylgalactosaminyltransferase

ST6GALNAC: alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase

UGT8: UDP Glycosyltransferase 8

PIGM: Phosphatidylinositol Glycan Anchor Biosynthesis Class M

STAT1: Signal Transducer and Activator of Transcription 1

HSF5: Heat Shock Transcription Factor 5

DLPFC: Dorsolateral Prefrontal Cortex

FP: Frontal Pole

IFG: Inferior Frontal Gyrus

STG: Superior Temporal Gyrus

PHG: Parahippocampal Gyrus

TCX: Temporal Cortex

CER: Cerebellar Cortex

MTC: Medial Temporal Cortex

LPFC: Lateral Prefrontal Cortex

LCBC: Lateral Cerebellar Cortex

TF: Transcription factors

TSS: Transcription start site

NES: Normalized Enrichment Score

Aβ: Amyloid β

APP: Amyloid Precursor Protein

MSBB: Mount Sinai Brain Bank

ROSMAP: Religious Orders Study/Memory and Aging Project

DECLARATIONS

Ethics approval and consent to participate

Study protocols were reviewed and approved by the Institutional Review Board (IRB). Written informed consent, including consent for autopsy, was obtained from study participants or, for those with substantial cognitive impairment, a caregiver, legal guardian or other proxy.

Availability of data and materials

The bulk RNA-seq datasets analyzed during the current study are available in the Accelerating Medicines Partnership Alzheimer's Disease Project (AMP-AD) Knowledge Portal under the RNAseq Harmonization Study (rnaSeqReprocessing), https://www.synapse.org/#!Synapse:syn17010685, but an access request is required. The N-glycomic datasets generated and analyzed are included in the supplementary information files of this article. Mass spectrometry data for glycomics (doi:10.25345/C5J67920T, MSV000090677) is available on MassIVE data repository upon reasonable request (contact: xctang@ucdavis.edu).

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the National Institute on Aging of the National Institutes of Health (R01AG062240) and University of California Davis Alzheimer's Disease Research Center (P30-AG072972, P30-AG010129). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contribution

Designed the study: AMZ, CBL, LWJ; Supervised the study: AMZ, CBL, LWJ, DH, IM; Performed experiments: JT, JDL, IM; Analyzed data: XT, JDL; Writing, review, and editing: XT, AMZ, JT, JDL, IM, DH, CBL, LWJ. All authors agree with the submission and have no conflicts of interest to disclose.

ACKNOWLEDGEMENTS

The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org). Data generation was supported by the following NIH grants: P30AG10161, P30AG72975, R01AG158 19, R01AG17917, R01AG036836, U01AG46152, U01AG61356, U01AG046139, P50 AG016574, R01 AG032990, U01AG046139, R01AG018023, U01AG006576, U01AG006786, R01AG025711, R01AG017216, R01AG003949, R01NS080820, U24NS072026, P30AG19610, U01AG046170, RF1AG057440, and U24AG061340, and the Cure PSP, Mayo and Michael J Fox foundations, Arizona Department of Health Services and the Arizona Biomedical Research Commission. We thank the participants of the Religious Order Study and Memory and Aging projects for the generous donation, the Sun Health Research Institute Brain and Body Donation Program, the Mayo Clinic Brain Bank, and the Mount Sinai/JJ Peters VA Medical Center NIH Brain and Tissue Repository. Data and analysis contributing investigators include Nilüfer Ertekin-Taner, Steven Younkin (Mayo Clinic, Jacksonville, FL), Todd Golde (University of Florida), Nathan Price (Institute for Systems Biology), David Bennett, Christopher Gaiteri (Rush University), Philip De Jager (Columbia University), Bin Zhang, Eric Schadt, Michelle Ehrlich, Vahram Haroutunian, Sam Gandy (Icahn School of Medicine at Mount Sinai), Koichi Iijima (National Center for Geriatrics and Gerontology, Japan), Scott Noggle (New York Stem Cell Foundation), Lara Mangravite (Sage Bionetworks). Special thanks to Dr. Chenghao Zhu for guidance and support in this project.

TABLES AND FIGURES

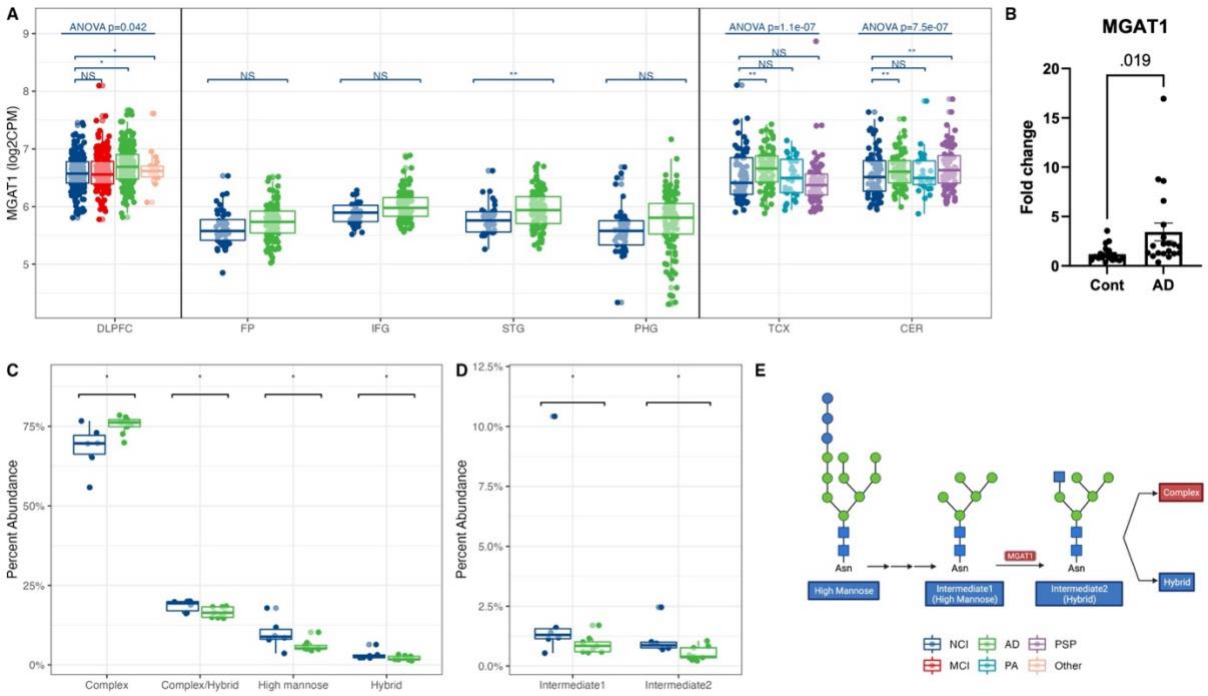


Figure 1-1: Increased MGAT1 expression corroborated by a raised complex and reduced high-mannose N-glycans abundance in AD. **A)** MGAT1 gene expression from RNA-seq across human brains with different diagnoses. Data presented as log2 count per million (log2CPM). Data from different studies were separated into blocks. Adjusted p-values were calculated based on all protein-coding genes. NS p-value >0.05, * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. ANOVA p-values were after the multiple testing correction over all protein-coding genes. DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia. **B)** qPCR confirmation of MGAT1 expression in the medial temporal cortex (MTC) of AD and control participants. **C)** The abundance of N-glycan

subtypes from N-glycomics in MTC of AD and control participants. **D)** The abundance of two intermediate structures (Hex:5 HexNAc: 2, Hex:5 HexNAc: 3) from N-glycomics in MTC of AD and control participants. *p-value<0.05. **E)** A schematic of the MGAT1 function. Red/blue boxes: higher/lower in AD than control. Glycan symbol key: green circles, mannose (Man); blue squares, N-acetylglucosamine (GlcNAc).

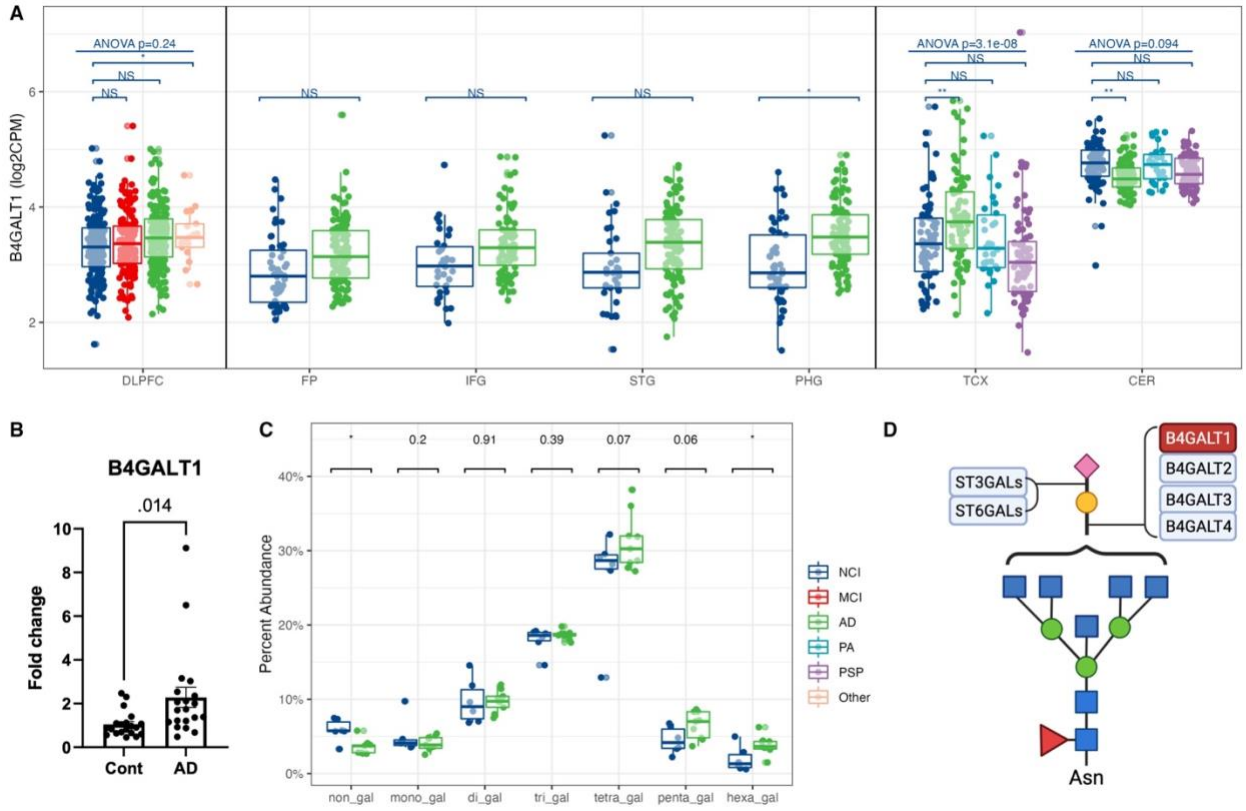


Figure 1-2: Increased B4GALT1 expression accompanied by a raised galactosylated N-glycans abundance in AD. **A)** B4GALT1 gene expression from RNA-seq across human brains with different diagnoses. Data presented as log2 count per million (log2CPM). Data from different studies were separated into blocks. Adjusted p-values were calculated based on all protein-coding genes. NS p-value >0.05, * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. ANOVA p-values were after the multiple testing correction over all protein-coding genes. DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia. **B)** qPCR confirmation of B4GALT1 expression in the medial temporal cortex (MTC) of AD and control participants. **C)** The abundance of complex and hybrid N-glycans with

different numbers of galactoses from N-glycomics in MTC of AD and control participants. *p-value<0.05. **D)** A schematic of galactosylation and sialylation steps. Red/blue boxes: over-/under-expressed in AD vs. control. Glycan symbol key: yellow circles, galactose (Gal); green circles, mannose (Man); blue squares, N-acetylglucosamine (GlcNAc); red triangles, fucose (Fuc); purple diamonds, N-acetylneuraminic acid (Neu5Ac).

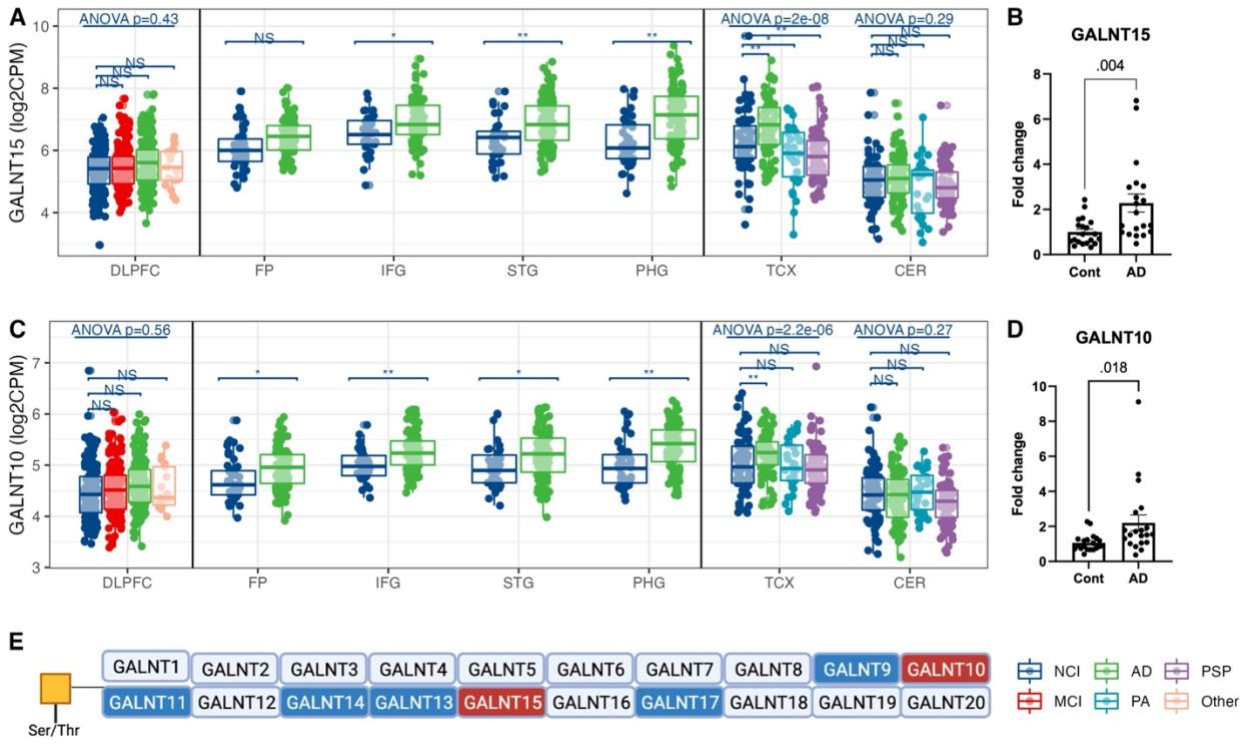


Figure 1-3: The differential expression of polypeptide N-acetylgalactosaminyltransferase (GALNT) gene family in AD across cortical regions. **A)** GALNT15 gene expression from RNA-seq across human brains with different diagnoses. Data presented as log2 count per million (log2CPM). Data from different studies were separated into blocks. Adjusted p-values were calculated based on all protein-coding genes. NS p-value >0.05, * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. ANOVA p-values were after the multiple testing correction over all protein-coding genes. DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia. **B)** qPCR confirmation of GALNT15 expression in the medial temporal cortex (MTC) of AD and control participants. **C)** GALNT10 gene expression from RNA-seq across human brains with different diagnoses. **D)** qPCR

confirmation of GALNT10 expression in MTC of AD and control participants. **E)** A schematic of the initiation step of O-GalNAc glycans. Red/blue boxes: over-/under-expressed in AD vs. control. Glycan symbol key: yellow squares, N-Acetyl-galactosamine (GalNAc).

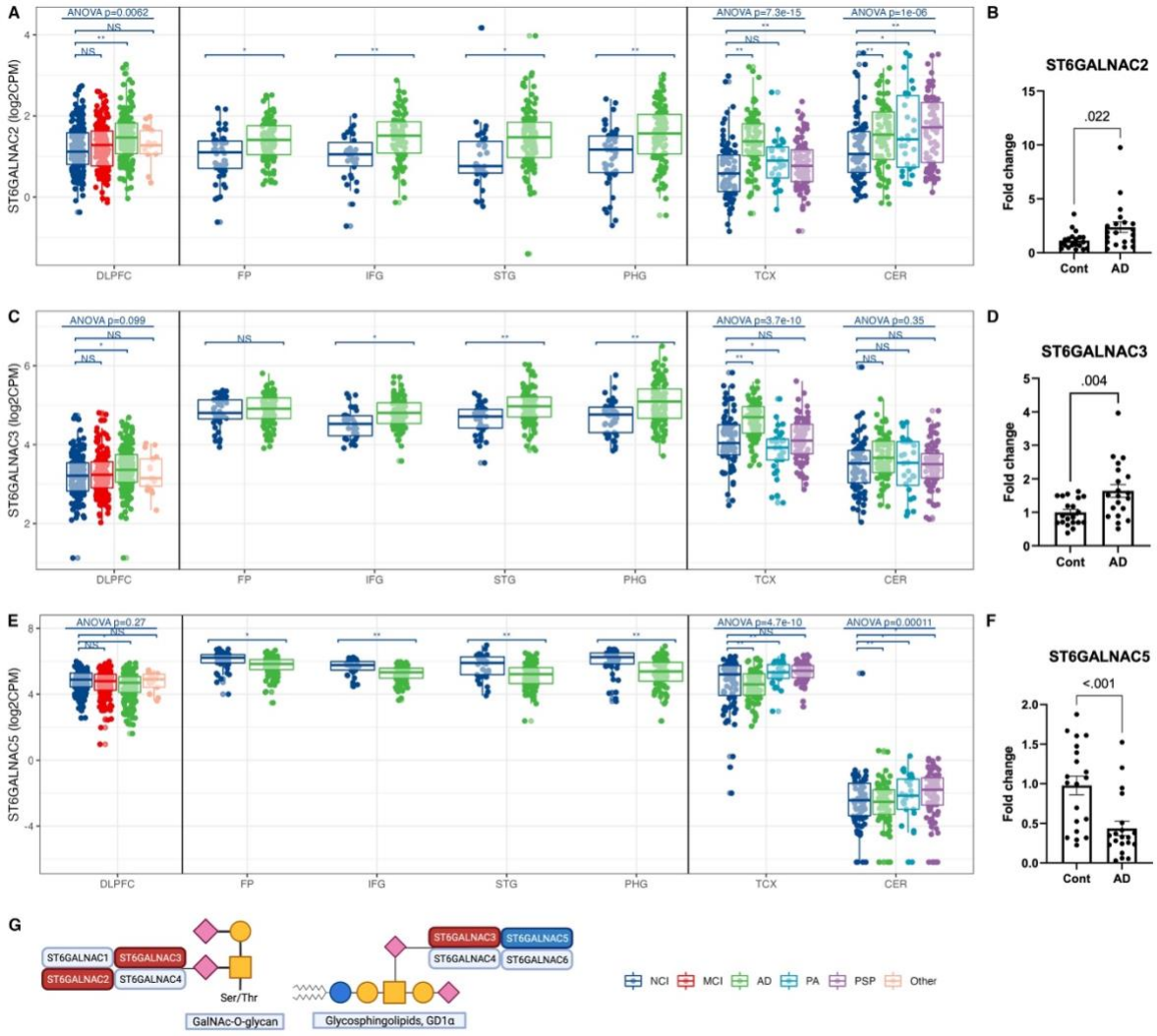


Figure 1-4: Differential expression of alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase (ST6GALNAC) gene family in AD across brain regions. **A)** ST6GALNAC2 gene expression from RNA-seq across human brains with different diagnoses. Data presented as log2 count per million (log2CPM). Data from different studies were separated into blocks. Adjusted p-values were calculated based on all protein-coding genes. NS p-value >0.05, * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. ANOVA p-values were after the multiple testing correction over all protein-coding genes. DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG

parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia. **B)** qPCR confirmation of ST6GALNAC2 expression in the medial temporal cortex (MTC) of AD and control participants. **C)** ST6GALNAC3 gene expression from RNA-seq across human brains with different diagnoses. **D)** qPCR confirmation of ST6GALNAC3 expression in the MTC of AD and control participants. **E)** ST6GALNAC5 gene expression from RNA-seq across human brains with different diagnoses. **F)** qPCR confirmation of ST6GALNAC5 expression in the MTC of AD and control participants. G) A schematic showing the function and specificity of ST6GALNACs. Red/blue boxes: over-/under-expressed in AD vs. control. Glycan symbol key: yellow circles, galactose (Gal); yellow squares, N-acetylgalactosamine (GalNAc); blue circles, N-glucose (Glc); purple diamonds, N-acetylneuraminic acid (Neu5Ac).

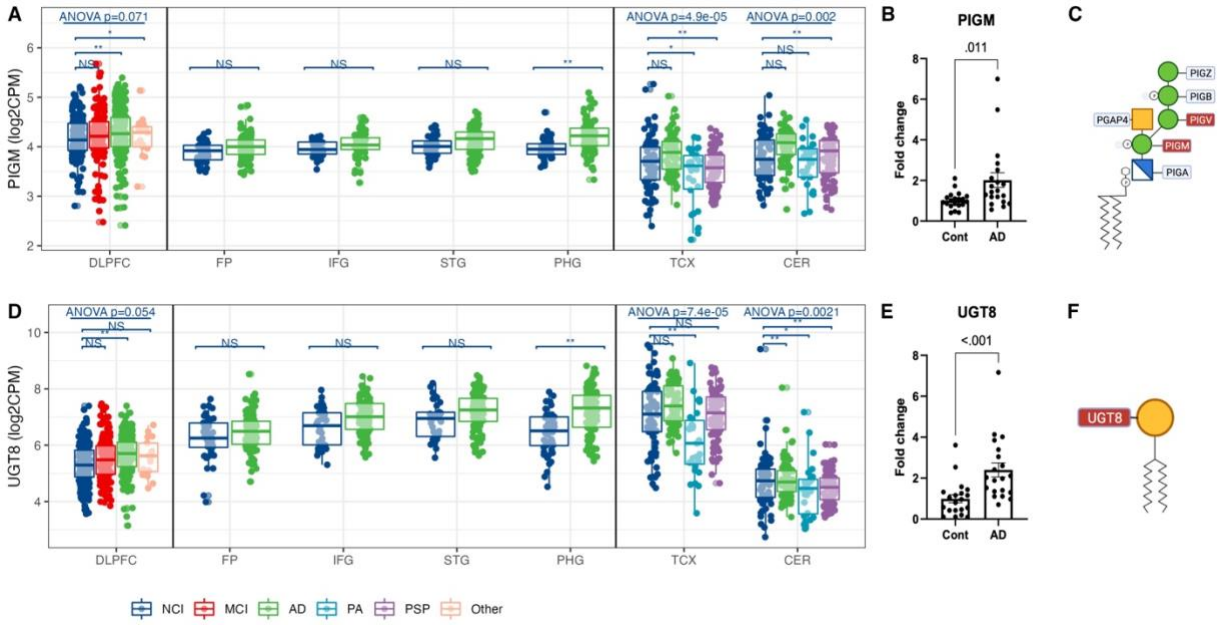


Figure 1-5: Differential expression of glycosyltransferases involved in GPI-anchor and glycosphingolipid biosynthesis in AD. **A)** PIGM gene expression from RNA-seq across human brains with different diagnoses. Data presented as log2 count per million (log2CPM). Data from different studies were separated into blocks. Adjusted p-values were calculated based on all protein-coding genes. NS p-value >0.05, * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. ANOVA p-values were after the multiple testing correction over all protein-coding genes. DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia. **B)** qPCR confirmation of PIGM expression in the medial temporal cortex (MTC) of AD and control participants. **C)** The schema of the GPI-anchor biosynthesis. Red/blue boxes: over-/under-expressed in AD vs. control. **D)** UGT8 gene expression from RNA-seq across human brains with different diagnoses. **E)** qPCR confirmation

of UGT8 expression in MTC of AD and control participants. **F)** A schematic of galactocerebroside biosynthesis. Red/blue boxes: over-/under-expressed in AD vs. control. Glycan symbol key: yellow circles, galactose (Gal); yellow squares, N-acetylgalactosamine (GalNAc); green circles, mannose (Man); blue cross squares, glucosamine (GlcN).

Supplemental Table 1-S1: Samples size information.

AD Study	Tissue	Diagnosis						Total Samples
		NCI	MCI	AD	PA	PSP	Other	
ROSMAP	DLPFC	195	161	208	-	-	17	581
MSBB	FP	48	-	114	-	-	-	162
	IFG	34	-	91	-	-	-	125
	STG	37	-	115	-	-	-	152
	PHG	49	-	110	-	-	-	159
Mayo	TCX	78	-	82	30	84	-	274
	CER	77	-	82	28	84	-	271
Total		518	161	802	58	168	17	1724

Notes: DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia.

Supplemental Table 1-S2: Demographic characteristics of the ROSMAP, MSBB, and Mayo Alzheimer's disease cohorts.

ROSMAP: DLPFC	**NCI**	**MCI**	**AD**	**Other**
N	195	161	208	17
Age (years) (median [IQR])	86.11[81.84-90]	89.42[85.38-90]	90[87.64-90]	88.96[85.90-90]
Sex=Male (%)	77(39.49%)	55(34.16%)	65(30.29%)	10(58.82%)
Edu (years) (median [IQR])	16[13.5-19]	16[14-18]	16[14-18]	18[15-21]
APOE genotype				
e22 (%)	2(1.03%)	1(0.62%)	1(0.48%)	0(0%)
e23 (%)	34(17.44%)	14(8.75%)	21(10.1%)	3(17.65%)
e24 (%)	2(1.03%)	4(2.5%)	8(3.85%)	1(5.88%)
e33 (%)	126(64.62%)	114(71.25%)	108(51.92%)	9(52.94%)
e34 (%)	30(15.38%)	27(16.88%)	68(32.69%)	3(17.65%)
e44 (%)	1(0.51%)	0(0%)	2(0.96%)	1(5.88%)
Ethnicity				
American Indian or Alaska Native (%)	0(0%)	1(0.62%)	0(0%)	0(0%)
Black (%)	2(1.03%)	1(0.62%)	2(0.96%)	0(0%)
White (%)	193(98.97%)	159(98.76%)	206(99.04%)	17(100%)
MMSE score (median [IQR])	28.75[27-29.5]	26[24-28]	15[5-20]	16[11-20]
MSBB: FP	**NCI**	**AD**		
N	48	114		
Age (years) (median [IQR])	82[74-89.25]	86[80-90]		
Sex=Male (%)	23(47.92%)	39(34.21%)		
APOE genotype				
e22 (%)	1(4.17%)	0(0%)		
e23 (%)	5(20.83%)	2(2.94%)		
e24 (%)	0(0%)	1(1.47%)		
e33 (%)	14(58.33%)	37(54.41%)		
e34 (%)	4(16.67%)	25(36.76%)		
e44 (%)	0(0%)	3(4.41%)		
Ethnicity				
American Native or Alaskan Native (%)	1(2.08%)	0(0%)		
Black (%)	6(12.5%)	10(8.77%)		
Hispanic (%)	6(12.5%)	4(3.51%)		

White (%)	35(72.92%)	99(86.84%)
Unknown (%)	0(0%)	1(0.88%)
CDR score (median [IQR])	0.25[0-0.5]	3[2-4]
MSBB: IFG	**NCI**	**AD**
N	34	91
Age (years) (median [IQR])	82.5[74.25-90]	86[80-90]
Sex=Male (%)	17(50%)	30(32.97%)
APOE genotype		
e22 (%)	0(0%)	0(0%)
e23 (%)	4(30.77%)	1(1.89%)
e24 (%)	0(0%)	1(1.89%)
e33 (%)	8(61.54%)	33(62.26%)
e34 (%)	1(7.69%)	15(28.3%)
e44 (%)	0(0%)	3(5.66%)
Ethnicity		
American Native or Alaskan Native (%)	0(0%)	0(0%)
Black (%)	4(11.76%)	8(8.79%)
Hispanic (%)	6(17.65%)	2(2.2%)
White (%)	24(70.59%)	80(87.91%)
Unknown (%)	0(0%)	1(1.1%)
CDR score (median [IQR])	0.5[0-0.5]	3[2-4]
MSBB: STG	**NCI**	**AD**
N	34	91
Age (years) (median [IQR])	82[74-89]	85[78.5-90]
Sex=Male (%)	16(43.24%)	41(35.65%)
APOE genotype		
e22 (%)	0(0%)	0(0%)
e23 (%)	5(27.78%)	4(6.35%)
e24 (%)	0(0%)	1(1.59%)
e33 (%)	10(55.56%)	36(57.14%)
e34 (%)	3(16.67%)	20(31.75%)
e44 (%)	0(0%)	2(3.17%)
Ethnicity		
American Native or Alaskan Native (%)	0(0%)	0(0%)
Black (%)	6(16.22%)	13(11.3%)
Hispanic (%)	6(16.22%)	3(2.61%)

White (%)	25(67.57%)	98(85.22%)		
Unknown (%)	0(0%)	1(0.87%)		
CDR score (median [IQR])	0[0-0.5]	3[3-5]		
MSBB: PHG	**NCI**	**AD**		
N	49	110		
Age (years) (median [IQR])	79[74-89]	85[80-90]		
Sex=Male (%)	24(48.98%)	40(36.36%)		
APOE genotype				
e22 (%)	1(3.57%)	0(0%)		
e23 (%)	7(25%)	3(5.88%)		
e24 (%)	0(0%)	1(1.96%)		
e33 (%)	16(57.14%)	30(58.82%)		
e34 (%)	4(14.29%)	14(27.45%)		
e44 (%)	0(0%)	3(5.88%)		
Ethnicity				
American Native or Alaskan Native (%)	1(2.04%)	0(0%)		
Black (%)	8(16.33%)	11(10%)		
Hispanic (%)	8(16.33%)	2(1.82%)		
White (%)	32(65.31%)	95(86.36%)		
Unknown (%)	0(0%)	2(1.82%)		
CDR score (median [IQR])	0[0-0.5]	3[2-5]		
Mayo: TCX	**NCI**	**AD**	**PA**	**PSP**
N	78	82	30	84
Age (years) (median [IQR])	86[79-89.75]	85[78.25-89]	86.5[83.25-88]	74[69-79]
Sex=Male (%)	41(52.56%)	33(40.24%)	13(43.33%)	51(60.71%)
APOE genotype				
e22 (%)	0(0%)	0(0%)	1(3.33%)	0(0%)
e23 (%)	12(15.38%)	4(4.88%)	3(10%)	15(17.86%)
e24 (%)	1(1.28%)	0(0%)	0(0%)	1(1.19%)
e33 (%)	57(73.08%)	35(42.68%)	16(53.33%)	56(66.67%)
e34 (%)	8(10.26%)	36(43.9%)	10(33.33%)	11(13.1%)
e44 (%)	0(0%)	7(8.54%)	0(0%)	1(1.19%)
Ethnicity				
White (%)	78 (%)	82(%)	30(%)	84(%)
Mayo: CER	**NCI**	**AD**	**PA**	**PSP**
N	77	82	28	84

Age (years) (median [IQR])	85[79-89]	84.5[78.25-89]	86[82.25-88]	74[69-79]
Sex=Male (%)	40(51.95%)	34(41.46%)	12(42.86%)	51(60.71%)
APOE genotype				
e22 (%)	0(0%)	0(0%)	1(3.57%)	0(0%)
e23 (%)	10(12.99%)	4(4.88%)	3(10.71%)	14(16.67%)
e24 (%)	1(1.3%)	0(0%)	0(0%)	1(1.19%)
e33 (%)	57(74.03%)	35(42.68%)	15(53.57%)	57(67.86%)
e34 (%)	9(11.69%)	38(46.34%)	9(32.14%)	11(13.1%)
e44 (%)	0(0%)	5(6.1%)	0(0%)	1(1.19%)
Ethnicity				
White (%)	77 (100%)	82 (100%)	28 (100%)	84 (100%)

Notes: Participants aged >90 were truncated at age 90 in this table. DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia.

Supplemental Table 1-S3: qPCR primer information.

Gene Name	Sequence
B4GALT1	FW: CTATATCTCGCCCAAATGCTG RV: GTGCAATTCGGTCAAACCTC
B3GALT5	FW: ATCAGGCAGCCATTCAGCAA RV: ACGTCGCCAGAAAACACGTA
GALNT15	FW: TGTCCTGAGGGTGCTGCT RV: AAGAGGAGTTGACCCCGTTC
GALNT11	FW: TGTCCAGGTCGCAGATGTAA RV: GTCGGACTGGGGCTGTG
GALNT10	FW: GCCCATCTCCCAAGAAAAA RV: GCTTTGGGCGCTGTACC
b-Actin	FW: TCAAGATCATTGCTCCTCCTGAG RV: ACATCTGCTGGAAGGTGGACA

Gene Name	Unique Assay ID
ST6GALNAC5	qHsaCID0007142
ST3GAL3	qHsaCED0036502
ALG1	qHsaCID0016546
ST6GALNAC3	qHsaCID0036897
ST3GAL5	qHsaCID0013823
ST6GALNAC2	qHsaCID0016785

GALNTL6	qHsaCED0057234
B3GALT2	qHsaCID0023450
NEU3	qHsaCED0045967
ST6GAL2	qHsaCID0007061
UGT8	qHsaCID0017816
B4GALT6	qHsaCID0015245
B3GALNT1	qHsaCED0046621
A4GALT	qHsaCED0002403
PIGM	qHsaCED0018886
B4GALT5	qHsaCID0015112

Supplemental Table 1-S4: Demographic data of participants in the glycomic analysis.

	LCBC		LPFC		MTC	
	AD	Control	AD	Control	AD	Control
n	7	10	10	9	9	6
Age (years) (median [IQR])	80[77.5-86]	85.5[79.75-90]	81[77.25-90]	87[79-90]	80[77-90]	88.5[84.75-90]
Sex=Male (%)	2(29)	3(30)	3(30)	3(33)	3(33)	2(33)
Ethnicity						
Black (%)	3(43)	1(10)	3(30)	1(11)	3(33)	1(17)
White (%)	4(57)	9(90)	7(70)	8(89)	6(67)	5(83)

Notes: LCBC lateral cerebellum cortex, LPFC lateral prefrontal cortex, MTC medial temporal cortex. Participants aged >90 were truncated at age 90 in this table.

Supplemental Table 1-S5: Glycosyltransferases excluded for low expression.

DLPFC	MGAT2/MGAT4D/GALNT4/GALNT5/LARGE2/B3GNT3/A4GNT/B4GALNT2/FUT5/FUT6/CHST4/HS3ST6/NDST4/NEU2
FP	MGAT4D/GALNT5/B3GNT6/LARGE2/B3GNT3/A4GNT/B4GALNT2/FUT3/FUT5/FUT6/FUT7/CHST13/HS3ST6/NEU2
STG	MGAT2/MGAT4D/GCNT3/B3GNT6/LARGE2/B3GNT3/A4GNT/B4GALNT2/FUT3/FUT5/FUT6/CHST13/HS3ST3A1/HS3ST6/NEU2
PHG	MGAT4D/GALNT4/GALNT5/GCNT3/B3GNT6/LARGE2/B3GNT3/A4GNT/B4GALNT2/FUT3/FUT5/FUT6/ST8SIA2/CHST13/HS3ST3A1/HS3ST6/NEU2
IFG	MGAT4D/GALNT5/GCNT3/B3GNT6/LARGE2/B3GNT3/A4GNT/B4GALNT2/FUT3/FUT5/FUT6/CHST13/HS3ST6/NEU2
CER	MGAT2/MGAT4D/GALNT4/GALNTL5 (T20)/B3GNT6/B3GNT3/A4GNT/B4GALNT2/FUT5/FUT6/CHST13/CHST4/HS3ST3A1/HS3ST6/NDST4/NEU2
TCX	MGAT2/MGAT4D/GALNT5/B3GNT6/LARGE2/B3GNT3/A4GNT/B4GALNT2/FUT5/FUT6/CHST4/HS3ST6/NEU2

Note: The gene expression cutoff is a minimum CPM of 10 in at least 70% of samples.

Supplemental Table 1-S6: Potential transcription factors and their target genes.

See the PDF file: SupplementaryTableS6TF_motif.pdf.

geneSet: The glycosylation pathway to which the glycosyltransferes belong.

Motif: ID of the motif.

NES: Normalized enrichment score of the motif in the gene-set. It was calculated based on the AUC distribution of all the motifs for the gene-set [(x-mean)/sd].

AUC: area under the curve.

TF.highConf: Transcription factors with high confident.

nEnrGenes: The number of genes that are highly ranked for the given motif.

enrichedGenes: hgnc_symbol of genes that are highly ranked for the given motif.

Supplemental Table 1-S7: miRNAs targeting the genes of glycosyltransferases according to validated databases (miRecords, miRTarBase and TarBase).

See the excel file: SupplementaryTableS7miRNA-targets.xlsx.

A: mature miRNAs targeting glycosyltransferases (GTs) for lipid glycosylation

B: mature miRNAs targeting glycosyltransferases (GTs) for N-glycosylation

C: mature miRNAs targeting glycosyltransferases (GTs) for mucin O-glycosylation

D: mature miRNAs targeting glycosyltransferases (GTs) for non-mucin O-glycosylation

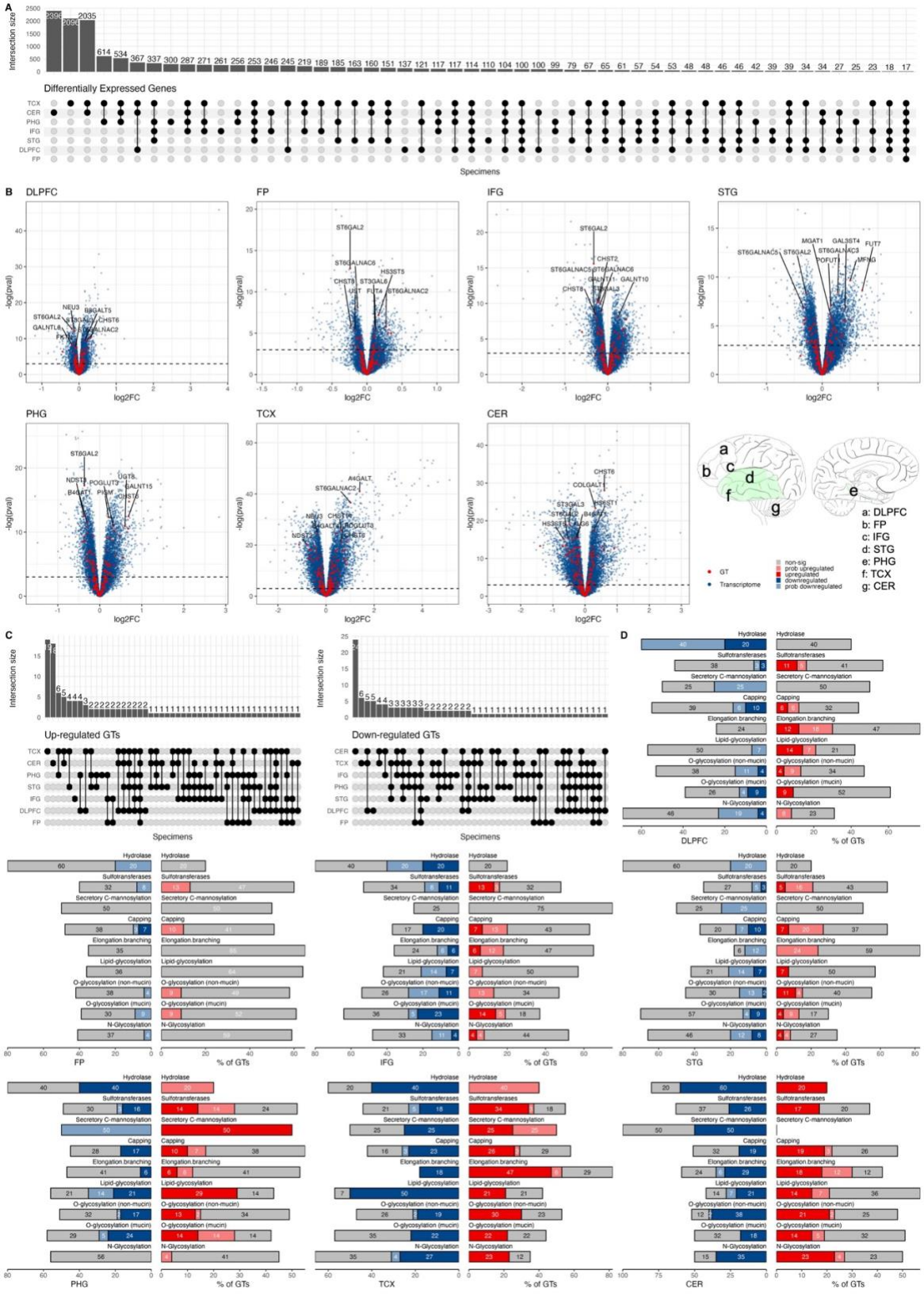
E: mature miRNAs targeting glycosyltransferases (GTs) for Secretory C-mannosylation

F: mature miRNAs targeting glycosyltransferases (GTs) for elongation and branching

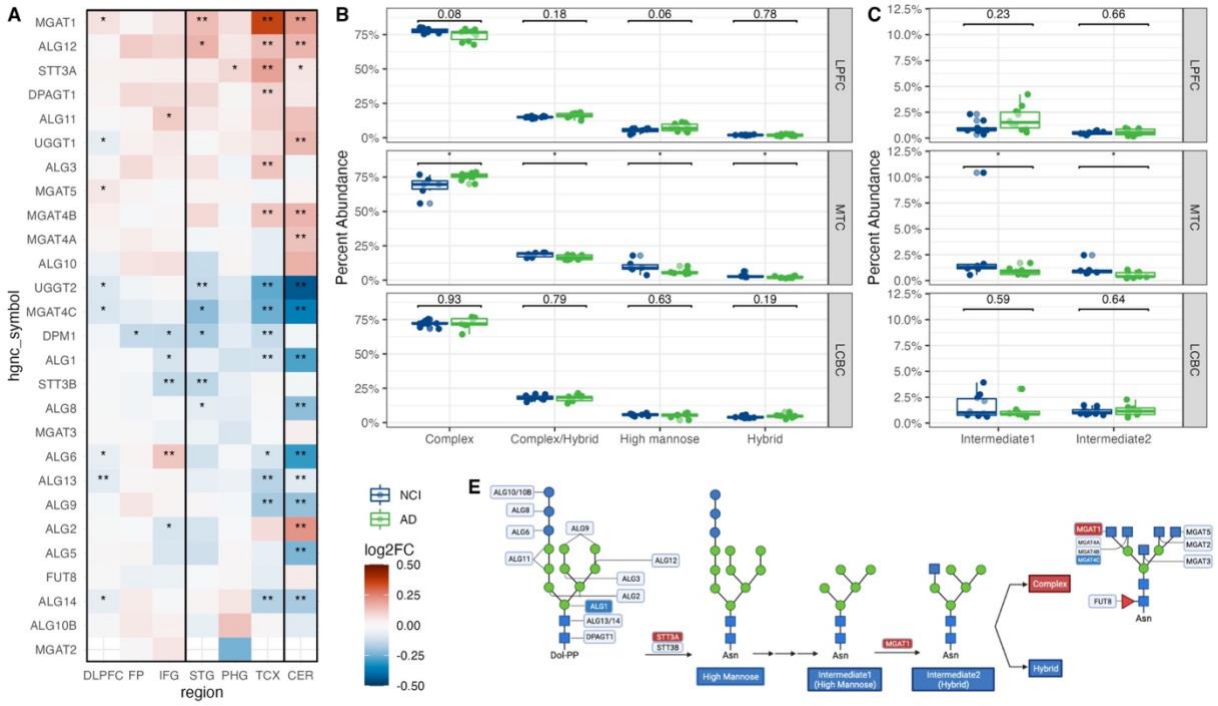
G: mature miRNAs targeting glycosyltransferases (GTs) for capping

H: mature miRNAs targeting glycosyltransferases (GTs) for sulfation

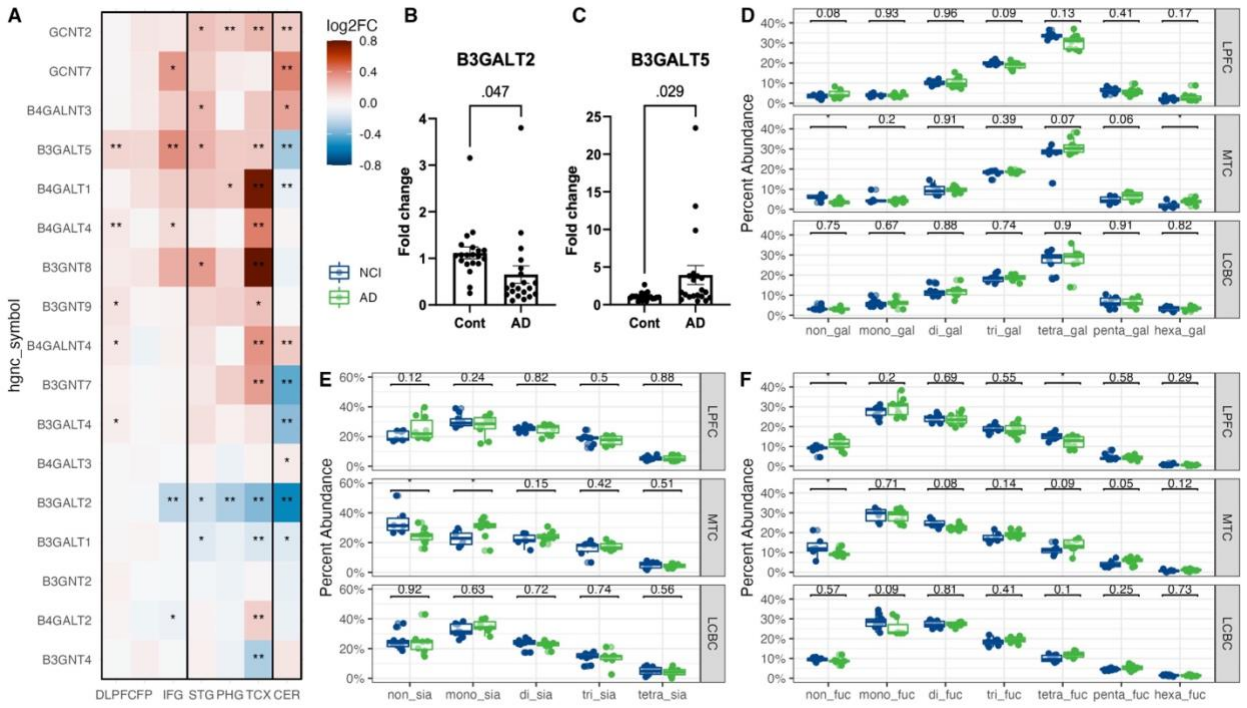
I: mature miRNAs targeting glycosyltransferases (GTs) for hydrolases


A
Intersection size
2500
2000
1500
1000
500
0
395
2094
2035
614 534 367 337 300 287 271 261 256 253 246 245 219 189 185 163 160 151 137 121 117 117 114 110 104 100 100 99 79 67 65 61 57 54 53 48 48 46 46 42 39 39 34 34 27 25 23 18 17
Differentially Expressed Genes
TCX
CER
PHG
IFG
STG
DLPFC
FP
Specimens
B
DLPFC
FP
IFG
STG
PHG
TCX
CER
-log(pval)
log2FC
ST6GAL2
NEU3
B3GALT5
ST3GAL6
CHST6
GALNTL6
FUT4
ST6GALNAC2
ST6GALNAC6
HS3ST5
CHST9
ST3GAL6
UST
FUT4
ST6GALNAC2
CHST2
ST6GALNAC5
ST6GALNAC6
GALNT10
CHST8
ST3GAL3
MGAT1
GAL3ST4
FUT7
ST6GALNAC5
ST6GALNAC2
ST6GALNAC3
POFUT1
MFNG
UGT8
NDST3
POGLUT3
GALNT15
B4GALT1
PIGM
CHST3
A4GALT
ST6GALNAC2
NEU3
CHST10
B4GALT1
POGLUT3
NDST3
GALNT3
CHST6
COL6GALT1
ST3GAL3
HS3ST1
ST6GAL1
B4GALT1
HS3ST1
NDST3
a
b
c
d
e
f
g
a: DLPFC
b: FP
c: IFG
d: STG
e: PHG
f: TCX
g: CER
GT
Transcriptome
non-sig
prob upregulated
upregulated
downregulated
prob downregulated
C
Intersection size
Up-regulated GTs
Down-regulated GTs
TCX
CER
PHG
STG
IFG
DLPFC
FP
Specimens
D
Hydrolase
Sulfotransferases
Secretory C-mannosylation
Capping
Elongation.branching
Lipid-glycosylation
O-glycosylation (non-mucin)
O-glycosylation (mucin)
N-Glycosylation
% of GTs
DLPFC
FP
IFG
STG
PHG
TCX
CER

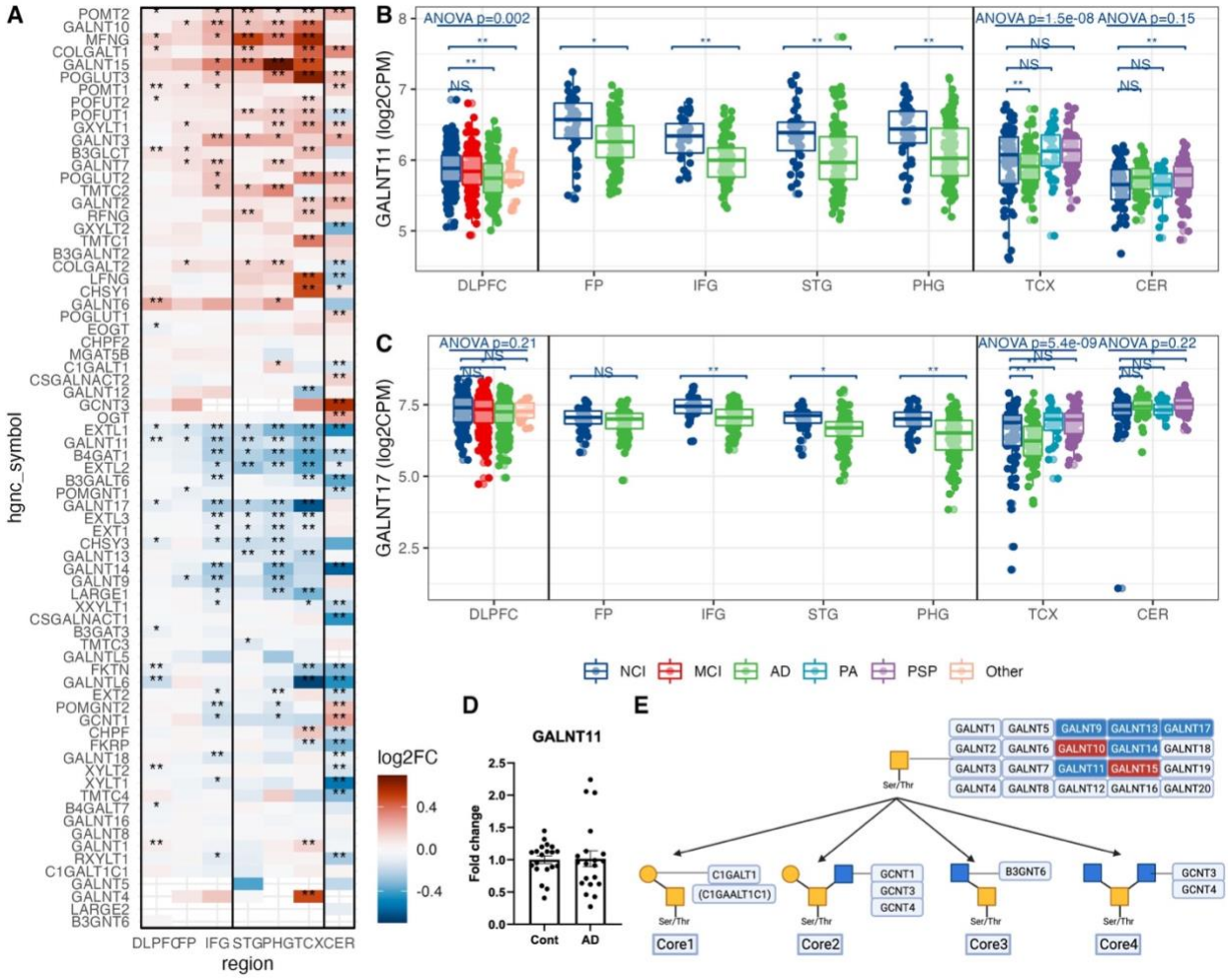
Supplemental Figure 1-S1: The overview of differentially expressed genes (DEGs) and glycosyltransferases (GTs) in AD vs. NCI using public RNA-seq data from MSBB, Mayo, and ROSMAP. A) The upset plot showing the number of DEGs overlapped between multiple regions (adjusted p-value < 0.05). B) Volcano plots of seven brain regions showing the distribution of genes and the magnitude of the differences. Red dots, glycosyltransferases (GTs); blue dots: protein-coding genes; the dashed line shows the significance cutoff (0.05). GTs with the lowest p-values were marked in the volcano plots. C) Upset plots showing the number of up-/down-regulated GTs overlapped between multiple regions (unadjusted p-value <0.05). The color showed the number of differentially expressed GTs. D) The percentage of GTs in each glycosylation pathway affect by AD in each brain region. Dark red/blue bars: GTs up-/down-regulated in AD vs. control with adjusted p-values<0.05; light red/blue bars: GTs probably upregulated in AD vs. control with p-values<0.05 but adjusted p-values>0.05; grey bars: GTs that didn't significantly changed in AD vs. control (p-values>0.05). Figures generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).



Supplemental Figure 1-S2: Differential expression of N-glycan genes and altered N-glycan abundances in brain specimens. A) The heatmap showed the glycosyltransferases (GTs) differential expression specific to N-glycosylation. Red/blue: up-/down-regulated in AD vs. NCI. * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. adjusted p-value were calculated based on all protein-coding genes. B) The abundance of N-glycan subtypes from N-glycomics in three brain regions of AD and control participants. C) The abundance of two intermediate structures (Hex:5 HexNAc: 2, Hex:5 HexNAc: 3) from N-glycomics in the three brain regions of AD and control participants. *p-value<0.05. LPFC: lateral prefrontal cortex; MTC: medial temporal cortex; LCBC: lateral cerebellar cortex. D) A schematic of the N-glycosylation process. Glycan symbol key: green circles, mannose (Man); blue squares, N-acetylglucosamine (GlcNAc); red triangles, fucose (Fuc). Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). The schematic diagram of glycans were created with BioRender.com.

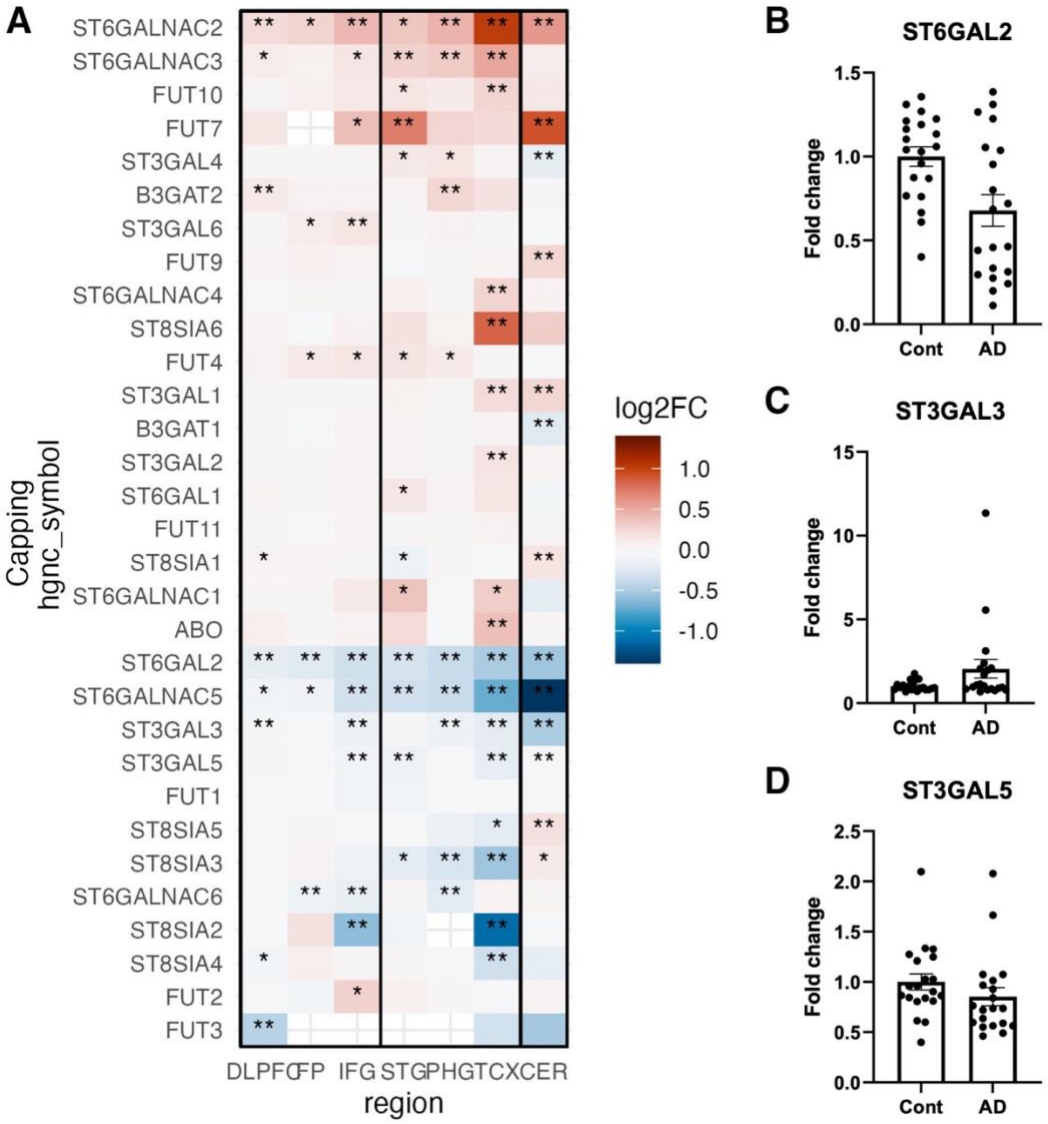


Supplemental Figure 1-S3: Differential expression of elongation genes and altered N-glycans abundances in brain specimens. A) The heatmap showed the glycosyltransferases (GTs) differential expression for elongation. Red/blue: up-/down-regulated in AD vs. NCI. * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. Adjusted p-values were calculated based on all protein-coding genes. B-C) The qPCR confirmation of B3GALT2 (B) and B3GALT5 (C). D) The abundance of N-glycans with different numbers of galactoses from N-glycomics in three brain regions of AD and control participants. E-F) The abundance of N-glycans with different numbers of Neu5Acs (E) and fucoses (F) from N-glycomics in three brain regions of AD and control participants. *p-value<0.05. LPFC: lateral prefrontal cortex; MTC: medial temporal cortex; LCBC: lateral cerebellar cortex. Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).

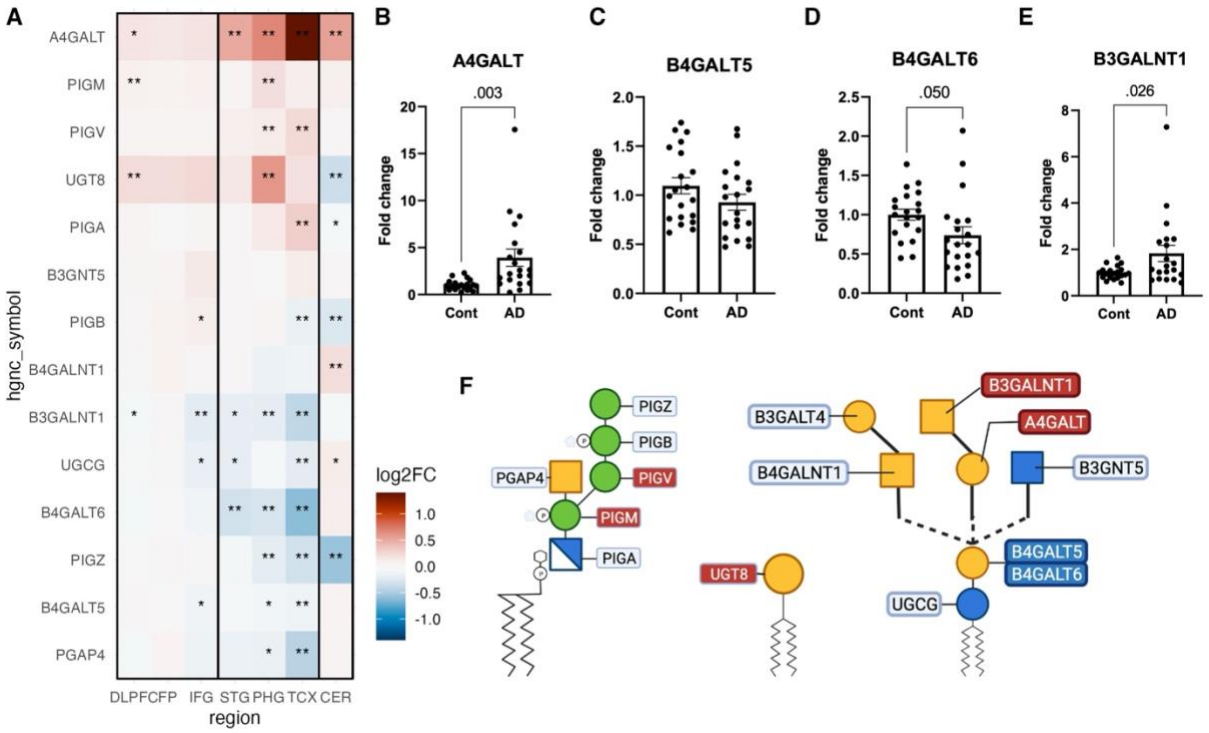


Supplemental Figure 1-S4: Differential expression of O-glycosylation genes in brain specimens. A) The heatmap showed the glycosyltransferases (GTs) differential expression for O-glycosylation. Red/blue: up-/down-regulated in AD vs. NCI. * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. Adjusted p-values were calculated based on all protein-coding genes. B-C) The changes of GALNT11 (B) and GALNT17 (C) across human brains with different diagnoses. Data presented as log2 count per million (log2CPM). Data from different studies were separated into blocks. Adjusted p-value were calculated based on all protein-coding genes. NS p-value >0.05, * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. ANOVA p-values were after the multiple testing correction over all protein-coding

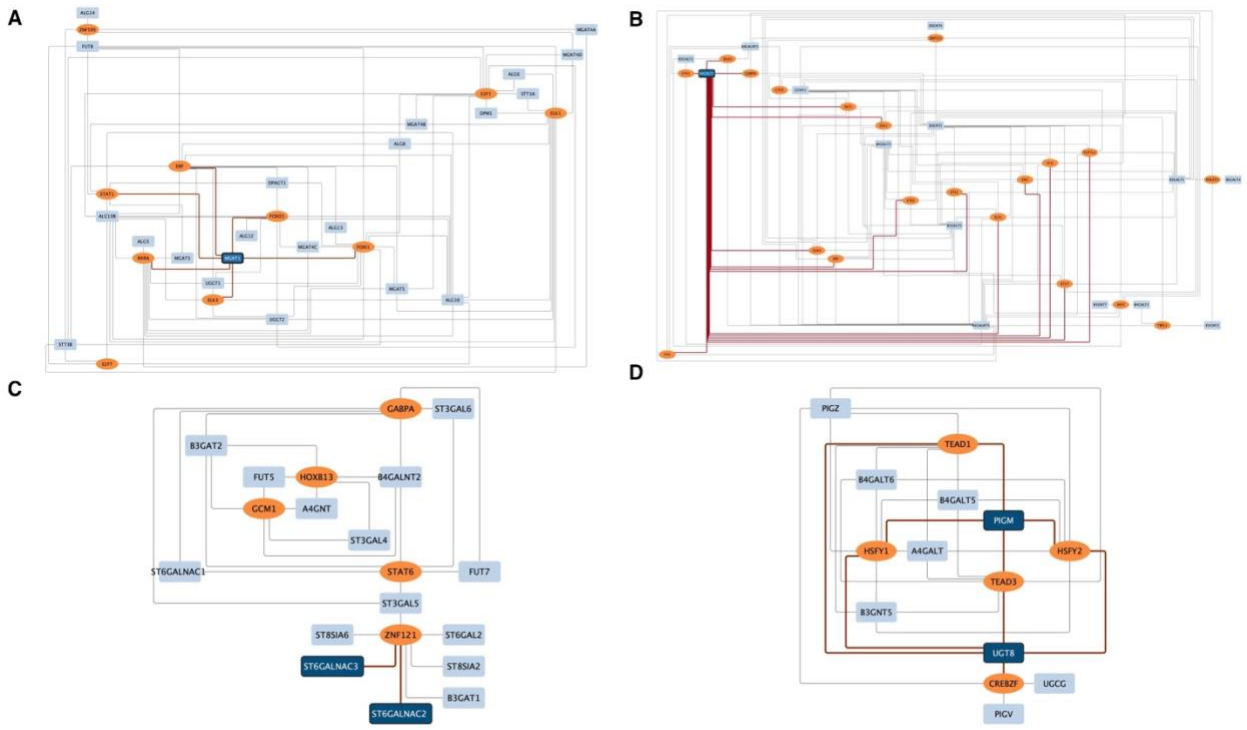
genes. DLPFC dorsolateral prefrontal cortex, FP frontal pole, IFG inferior frontal gyrus, STG superior temporal gyrus, PHG parahippocampal gyrus, TCX temporal cortex, CER cerebellum. NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease; PA: pathological aging; PSP: progressive supranuclear palsy; Other: other types of dementia. D) qPCR results of GALNT11 expression in MTC of AD and control participants. E) A schematic of the O-GalNAc-glycosylation process. Glycan symbol key: yellow circles, galactose (Gal); yellow squares, N-acetylgalactosamine (GalNAc); blue squares, N-acetylglucosamine (GlcNAc). Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). The schematic diagram of glycans were created with BioRender.com.



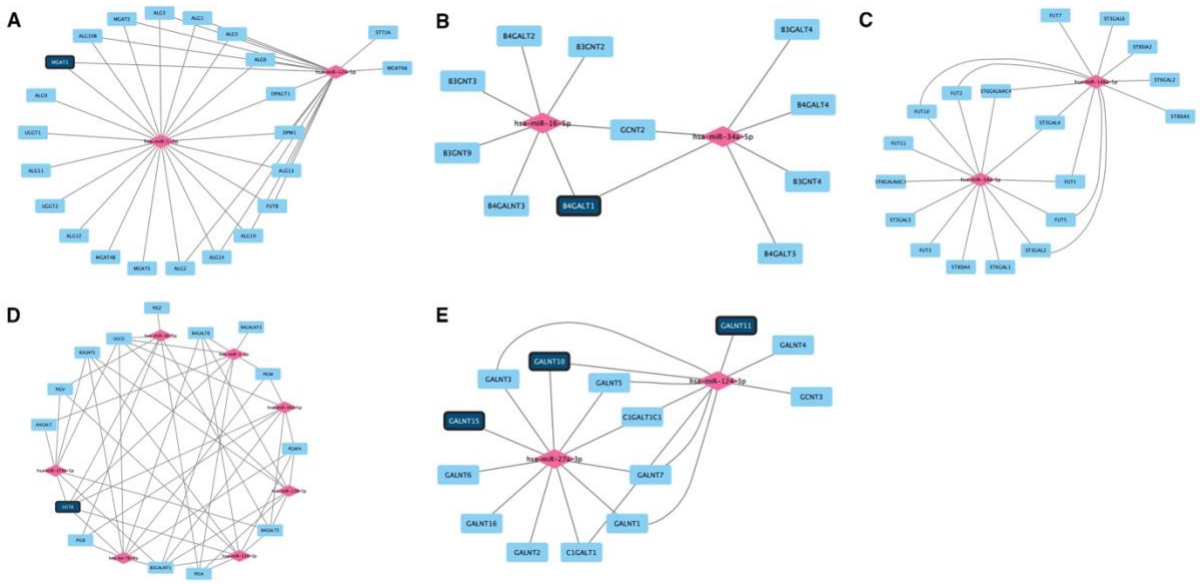
Supplemental Figure 1-S5: Differential expression of capping genes in brain specimens. A) The heatmap showed the glycosyltransferases (GTs) differential expression for capping. Red/blue: up-/down-regulated in AD vs. NCI. * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. Adjusted p-values were calculated based on all protein-coding genes. B-D) qPCR results of ST6GAL2 (B), ST3GAL3 (C), and ST3GAL5 (D) expression in MTC of AD and control participants. Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).



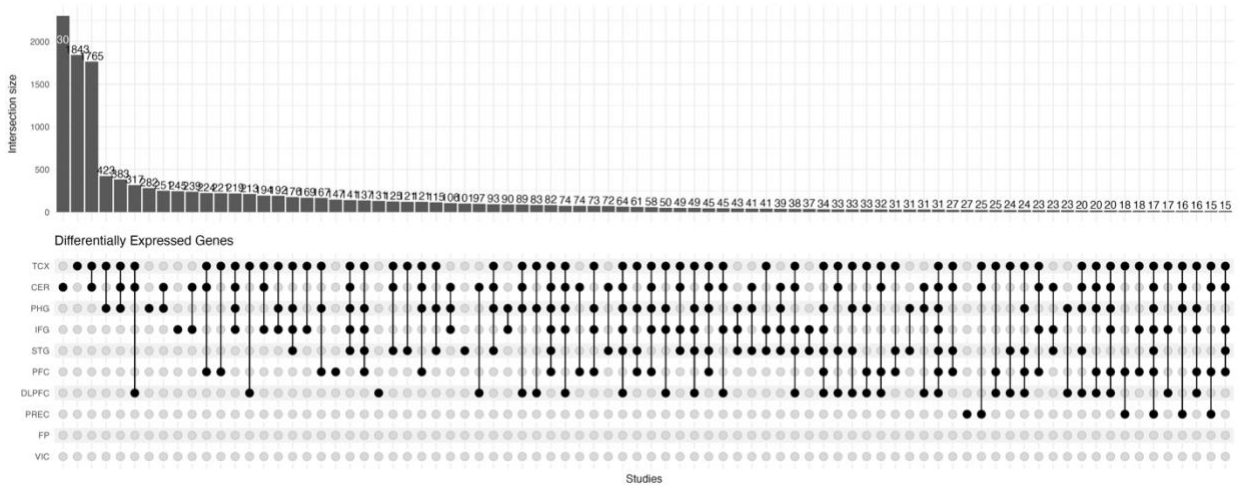
Supplemental Figure 1-S6: Differential expression of lipid glycosylation genes in brain specimens. A) The heatmap showed the glycosyltransferases (GTs) differential expression for lipid glycosylation. Red/blue: up-/down-regulated in AD vs. NCI. * p-value <0.05 but adjusted p-value>0.05, ** p-value<0.05 and adjusted p-value<0.05. Adjusted p-values were calculated based on all protein-coding genes. B-E) qPCR results of A4GALT (B), B4GALT5 (C), B4GALT6 (D), and B3GALNT1 (E) expression in MTC of AD and control participants. F) A schematic of the GPI-anchor, galactosylceramides, and glycosphingolipids biosynthesis. Glycan symbol key: yellow circles, galactose (Gal); yellow squares, N-acetylgalactosamine (GalNAc); blue circles, N-glucose (Glc); blue squares, N-acetylglucosamine (GlcNAc); green circles, mannose (Man); blue cross squares, glucosamine (GlcN). Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). The schematic diagram of glycans were created with BioRender.com.



Supplemental Figure 1-S7: Predicted transcription factors (TF) responsible for glycosylation. Regulatory networks for A) N-glycan core structure biosynthesis, B) elongation and branching, C) capping, and D) glycolipid biosynthesis. TFs with normalized enrichment score (NES) > 4.5 were shown in the network. TFs were represented as orange ellipse; glycosyltransferases were represented as blue rectangles. The significant glycosyltransferases were highlighted with in a dark shade of blue. The brown edges showed interactions between significant genes and TFs. Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).



Supplemental Figure 1-S8: Top 2 miRNA regulated the most number of glycosyltransferases and their validated targets. Regulatory networks for A) N-glycan core structure biosynthesis, B) elongation and branching, C) capping, and D) glycolipid biosynthesis, and E) O-GalNAc-glycan core structure biosynthesis. miRNAs were represented as purple diamonds; glycosyltransferases were represented as blue rectangles. The significant glycosyltransferases were highlighted with in a dark shade of blue. The brown edges showed interactions between significant genes and miRNAs. Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).



Supplemental Figure 1-S9: The upset plot showing the number of DEGs overlapped between our study (adjusted p-value < 0.05) and other publications [10,11]. Figures were generated using R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria).

REFERENCES

[1] Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's Dement 2018;14:535–62. https://doi.org/10.1016/j.jalz.2018.02.018.

[2] Stelzmann RA, Norman Schnitzlein H, Reed Murtagh F. An english translation of alzheimer's 1907 paper, "über eine eigenartige erkankung der hirnrinde." Clin Anat 1995;8:429–31. https://doi.org/10.1002/ca.980080612.

[3] Pahnke J, Walker LC, Scheffler K, Krohn M. Alzheimer's disease and blood-brain barrier function-Why have anti-β-amyloid therapies failed to prevent dementia progression? Neurosci Biobehav Rev 2009;33:1099–108. https://doi.org/10.1016/j.neubiorev.2009.05.006.

[4] Soejitno A, Tjan A, Purwata TE. Alzheimer's Disease: Lessons Learned from Amyloidocentric Clinical Trials. CNS Drugs 2015;29:487–502. https://doi.org/10.1007/s40263-015-0257-8.

[5] Klarić TS, Lauc G. The dynamic brain N-glycome. Glycoconj J 2022;39:443–71. https://doi.org/10.1007/s10719-022-10055-x.

[6] Frenkel-Pinter M, Shmueli MD, Raz C, Yanku M, Zilberzwige S, Gazit E, et al. Interplay between protein glycosylation pathways in Alzheimer's disease. Sci Adv 2017;3. https://doi.org/10.1126/sciadv.1601576.

[7] Gizaw ST, Ohashi T, Tanaka M, Hinou H, Nishimura SI. Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. Biochim Biophys Acta - Gen Subj 2016;1860:1716–27. https://doi.org/10.1016/j.bbagen.2016.03.009.

[8] Sihlbom C, Davidsson P, Sjögren M, Wahlund LO, Nilsson CL. Structural and quantitative comparison of cerebrospinal fluid glycoproteins in Alzheimer's disease patients and healthy individuals. Neurochem Res 2008;33:1332–40. https://doi.org/10.1007/s11064-008-9588-x.

[9] Kanninen K, Goldsteins G, Auriola S, Alafuzoff I, Koistinaho J. Glycosylation changes in Alzheimer's disease as revealed by a proteomic approach. Neurosci Lett 2004;367:235–40. https://doi.org/10.1016/j.neulet.2004.06.013.

[10] Zhang Q, Ma C, Chin LS, Li L. Integrative glycoproteomics reveals protein n-glycosylation aberrations and glycoproteomic network alterations in Alzheimer's disease. Sci Adv 2020;6. https://doi.org/10.1126/sciadv.abc5802.

[11] Estus S, Shaw BC, Devanney N, Katsumata Y, Press EE, Fardo DW. Evaluation of CD33 as a genetic risk factor for Alzheimer's disease. Acta Neuropathol 2019;138:187–99. https://doi.org/10.1007/s00401-019-02000-4.

[12] Freeze HH, Eklund EA, Ng BG, Patterson MC. Neurological Aspects of Human Glycosylation Disorders. Annu Rev Neurosci 2015;38:105–25. https://doi.org/10.1146/annurev-neuro-071714-034019.

[13] Ye Z, Marth JD. N-glycan branching requirement in neuronal and postnatal viability. Glycobiology 2004;14:547–58. https://doi.org/10.1093/glycob/cwh069.

[14] Liu F, Xu K, Xu Z, De Las Rivas M, Wang C, Li X, et al. The small molecule luteolin inhibits N-acetyl-α-galactosaminyltransferases and reduces mucin-type O-glycosylation of amyloid precursor protein. J Biol Chem 2017;292:21304–19. https://doi.org/10.1074/jbc.M117.814202.

[15] Akasaka-Manya K, Kawamura M, Tsumoto H, Saito Y, Tachida Y, Kitazume S, et al. Excess APP O-glycosylation by GalNAc-T6 decreases Ab production. J Biochem 2017;161:99–111. https://doi.org/10.1093/jb/mvw056.

[16] Fang P, Xie J, Sang S, Zhang L, Liu M, Yang L, et al. Multilayered N-Glycoproteome Profiling Reveals Highly Heterogeneous and Dysregulated Protein N-Glycosylation Related to Alzheimer's Disease. Anal Chem 2020;92:867–74. https://doi.org/10.1021/acs.analchem.9b03555.

[17] McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 2012;40:4288–97. https://doi.org/10.1093/nar/gks042.

[18] Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. Nat Rev Mol Cell Biol 2020;21:729–49. https://doi.org/10.1038/s41580-020-00294-x.

[19] Montine TJ, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Dickson DW, et al. National institute on aging-Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease: A practical approach. Acta Neuropathol 2012;123:1–11. https://doi.org/10.1007/s00401-011-0910-3.

[20] Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: Single-cell regulatory network inference and clustering. Nat Methods 2017;14:1083–6. https://doi.org/10.1038/nmeth.4463.

[21] Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, et al. The multiMiR R package and database: Integration of microRNA-target interactions along with their disease and drug associations. Nucleic Acids Res 2014;42. https://doi.org/10.1093/nar/gku631.

[22] Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Softw 2017;82:1–26. https://doi.org/10.18637/JSS.V082.I13.

[23] Guennewig B, Lim J, Marshall L, McCorkindale AN, Paasila PJ, Patrick E, et al. Defining early changes in Alzheimer's disease from RNA sequencing of brain regions differentially affected by pathology. Sci Rep 2021;11. https://doi.org/10.1038/s41598-021-83872-z.

[24] Williams JB, Cao Q, Yan Z. Transcriptomic analysis of human brains with Alzheimer's disease reveals the altered expression of synaptic genes linked to cognitive deficits. Brain Commun 2021;3. https://doi.org/10.1093/braincomms/fcab123.

[25] Wan YW, Al-Ouran R, Mangleburg CG, Perumal TM, Lee T V., Allison K, et al. Meta-Analysis of the Alzheimer's Disease Human Brain Transcriptome and Functional Dissection in Mouse Models. Cell Rep 2020;32:107908. https://doi.org/10.1016/j.celrep.2020.107908.

[26] Cho BG, Veillon L, Mechref Y. N-Glycan Profile of Cerebrospinal Fluids from Alzheimer's Disease Patients Using Liquid Chromatography with Mass Spectrometry. J Proteome Res 2019;18:3770–9. https://doi.org/10.1021/acs.jproteome.9b00504.

[27] Ioffe E, Stanley P. Mice lacking N-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. Proc Natl Acad Sci U S A 1994;91:728–32. https://doi.org/10.1073/pnas.91.2.728.

[28] Gaunitz S, Tjernberg LO, Schedin-Weiss S. The N-glycan profile in cortex and hippocampus is altered in Alzheimer disease. J Neurochem 2021;159:292–304. https://doi.org/10.1111/jnc.15202.

[29] Williams SE, Noel M, Lehoux S, Cetinbas M, Xavier RJ, Sadreyev RI, et al. Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. Nat Commun 2022;13. https://doi.org/10.1038/s41467-021-27781-9.

[30] Niu S, Fei M, Cheng C, Yan M, Gao S, Chen M, et al. Altered β-1,4-galactosyltransferase I expression during early inflammation after spinal cord contusion injury. J Chem Neuroanat 2008;35:245–56. https://doi.org/10.1016/j.jchemneu.2008.01.002.

[31] Wang P, Li X, Xie Y. B4GalT1 Regulates Apoptosis and Autophagy of Glioblastoma In Vitro and In Vivo. Technol Cancer Res Treat 2020;19. https://doi.org/10.1177/1533033820980104.

[32] Narimatsu Y, Joshi HJ, Nason R, Van Coillie J, Karlsson R, Sun L, et al. An Atlas of Human Glycosylation Pathways Enables Display of the Human Glycome by Gene Engineered Cells. Mol Cell 2019;75:394-407.e5. https://doi.org/10.1016/j.molcel.2019.05.017.

[33] Stanley P. What Have We Learned from Glycosyltransferase Knockouts in Mice? J Mol Biol 2016;428:3166–82. https://doi.org/10.1016/j.jmb.2016.03.025.

[34] Huang YF, Aoki K, Akase S, Ishihara M, Liu YS, Yang G, et al. Global mapping of glycosylation pathways in human-derived cells. Dev Cell 2021;56:1195-1209.e7. https://doi.org/10.1016/j.devcel.2021.02.023.

[35] Dworkin LA, Clausen H, Joshi HJ. Applying transcriptomics to studyglycosylation at the cell type level. IScience 2022;25:104419. https://doi.org/10.1016/j.isci.2022.104419.

[36] Steentoft C, Vakhrushev SY, Vester-Christensen MB, Schjoldager KTBG, Kong Y, Bennett EP, et al. Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. Nat Methods 2011;8:977–82. https://doi.org/10.1038/nmeth.1731.

[37] de las Rivas M, Lira-Navarrete E, Gerken TA, Hurtado-Guerrero R. Polypeptide GalNAc-Ts: from redundancy to specificity. Curr Opin Struct Biol 2019;56:87–96. https://doi.org/10.1016/j.sbi.2018.12.007.

[38] Schjoldager KTBG, Vakhrushev SY, Kong Y, Steentoft C, Nudelman AS, Pedersen NB, et al. Probing isoform-specific functions of polypeptide GalNAc-transferases using zinc finger nuclease glycoengineered SimpleCells. Proc Natl Acad Sci U S A 2012;109:9893–8. https://doi.org/10.1073/pnas.1203563109.

[39] Cheng L, Tachibana K, Zhang Y, Guo J ming, Kahori Tachibana K, Kameyama A, et al. Characterization of a novel human UDP-GalNAc transferase, pp-GalNAc-T10. FEBS Lett 2002;531:115–21. https://doi.org/10.1016/S0014-5793(02)03399-9.

[40] Raman J, Guan Y, Perrine CL, Gerken TA, Tabak LA. UDP-N-acetyl α-d-galactosamine: Polypeptide N- acetylgalactosaminyltransferases: Completion of the family tree. Glycobiology 2012;22:768–77. https://doi.org/10.1093/glycob/cwr183.

[41] Lee YC, Kaufmann M, Kitazume-Kawaguchi S, Kono M, Takashima S, Kurosawa N, et al. Molecular cloning and functional expression of two members of mouse NeuAcα2,3Galβ1,3GalNac GalNacα2,6-sialyltransferase family, ST6GalNAc III and iv. J Biol Chem 1999;274:11958–67. https://doi.org/10.1074/jbc.274.17.11958.

[42] Kurosawa N, Kojima N, Inoue M, Hamamoto T, Tsuji S. Cloning and expression of Gal beta 1,3GalNAc-specific GalNAc alpha 2,6-sialyltransferase. J Biol Chem 1994;269:19048–53. https://doi.org/10.1016/s0021-9258(17)32272-x.

[43] Okajima T, Fukumoto S, Ito H, Kiso M, Hirabayashi Y, Urano T, et al. Molecular cloning of brain-specific GD1α synthase (ST6GalNAc V) containing CAG/glutamine repeats. J Biol Chem 1999;274:30557–62. https://doi.org/10.1074/jbc.274.43.30557.

[44] Wißfeld J, Nozaki I, Mathews M, Raschka T, Ebeling C, Hornung V, et al. Deletion of Alzheimer's disease-associated CD33 results in an inflammatory human microglia phenotype. Glia 2021;69:1393–412. https://doi.org/10.1002/glia.23968.

[45] Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat Genet 2011;43:436–43. https://doi.org/10.1038/ng.801.

[46] Malik M, Simpson JF, Parikh I, Wilfred BR, Fardo DW, Nelson PT, et al. CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. J Neurosci 2013;33:13320–5. https://doi.org/10.1523/JNEUROSCI.1224-13.2013.

[47] Gong L, Xu R, Lan L, Liu D, Shen J, Zhang B. The CD33 genotype associated cognitive performance was bidirectionally modulated by intrinsic functional connectivity in the Alzheimer's disease spectrum. Biomed Pharmacother 2019;115. https://doi.org/10.1016/j.biopha.2019.108903.

[48] Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 2019;51:404–13. https://doi.org/10.1038/s41588-018-0311-9.

[49] Varki A, Angata T. Siglecs - The major subfamily of I-type lectins. Glycobiology 2006;16. https://doi.org/10.1093/glycob/cwj008.

[50] Bhattacherjee A, Daskhan GC, Bains A, Watson AES, Eskandari-Sedighi G, St. Laurent CD, et al. Increasing phagocytosis of micoglia by targeting CD33 with liposomes displaying glycan ligands. J Control Release 2021;338:680–93. https://doi.org/10.1016/j.jconrel.2021.09.010.

[51] Rocha EM, Smith GA, Park E, Cao H, Brown E, Hallett P, et al. Progressive decline of glucocerebrosidase in aging and P arkinson's disease. Ann Clin Transl Neurol 2015;2:433–8. https://doi.org/10.1002/acn3.177.

[52] Belarbi K, Cuvelier E, Bonte MA, Desplanque M, Gressier B, Devos D, et al. Glycosphingolipids and neuroinflammation in Parkinson's disease. Mol Neurodegener 2020;15. https://doi.org/10.1186/s13024-020-00408-1.

[53] Sandhoff R, Sandhoff K. Emerging concepts of ganglioside metabolism. FEBS Lett 2018;592:3835–64. https://doi.org/10.1002/1873-3468.13114.

[54] NORTON WT, AUTILIO LA. THE LIPID COMPOSITION OF PURIFIED BOVINE BRAIN MYELIN. J Neurochem 1966;13:213–22. https://doi.org/10.1111/j.1471-4159.1966.tb06794.x.

[55] Tan RPA, Leshchyns'Ka I, Sytnyk V. Glycosylphosphatidylinositol-anchored immunoglobulin superfamily cell adhesion molecules and their role in neuronal development and synapse regulation. Front Mol Neurosci 2017;10. https://doi.org/10.3389/fnmol.2017.00378.

[56] Puzzo D, Bizzoca A, Privitera L, Furnari D, Giunta S, Girolamo F, et al. F3/Contactin promotes hippocampal neurogenesis, synaptic plasticity, and memory in adult mice. Hippocampus 2013;23:1367–82. https://doi.org/10.1002/hipo.22186.

[57] Rangaraju S, Dammer EB, Raza SA, Rathakrishnan P, Xiao H, Gao T, et al. Identification and therapeutic modulation of a pro-inflammatory subset of disease-associated-microglia in Alzheimer's disease. Mol Neurodegener 2018;13. https://doi.org/10.1186/s13024-018-0254-8.

[58] Butturini E, Boriero D, Carcereri de Prati A, Mariotto S. STAT1 drives M1 microglia activation and neuroinflammation under hypoxia. Arch Biochem Biophys 2019;669:22–30. https://doi.org/10.1016/j.abb.2019.05.011.

[59] Hsu WL, Ma YL, Hsieh DY, Liu YC, Lee EH. STAT1 negatively regulates spatial memory formation and mediates the memory-impairing effect of Aβ. Neuropsychopharmacology 2014;39:746–58. https://doi.org/10.1038/npp.2013.263.

[60] Liu YX, Wang L, Liu WJ, Zhang HT, Xue JH, Zhang ZW, et al. MIR-124-3p/B4GALT1 axis plays an important role in SOCS3-regulated growth and chemo-sensitivity of CML. J Hematol Oncol 2016;9. https://doi.org/10.1186/s13045-016-0300-3.

[61] Fang M, Wang J, Zhang X, Geng Y, Hu Z, Rudd JA, et al. The miR-124 regulates the expression of BACE1/β-secretase correlated with cell death in Alzheimer's disease. Toxicol Lett 2012;209:94–105. https://doi.org/10.1016/j.toxlet.2011.11.032.

[62] An F, Gong G, Wang Y, Bian M, Yu L, Wei C. MiR-124 acts as a target for Alzheimer's disease by regulating BACE1. Oncotarget 2017;8:114065–71. https://doi.org/10.18632/oncotarget.23119.

[63] Cai H, Zhou H, Miao Y, Li N, Zhao L, Jia L. MiRNA expression profiles reveal the involvement of miR-26a, miR-548l and miR-34a in hepatocellular carcinoma progression through regulation of ST3GAL5. Lab Investig 2017;97:530–42. https://doi.org/10.1038/labinvest.2017.12.

[64] Li W, Ma H, Sun J. microRNA-34a/c function as tumor suppressors in Hep-2 laryngeal carcinoma cells and may reduce GALNT7 expression. Mol Med Rep 2014;9:1293–8. https://doi.org/10.3892/mmr.2014.1929.

[65] Bernardi C, Soffientini U, Piacente F, Tonetti MG. Effects of MicroRNAs on Fucosyltransferase 8 (FUT8) Expression in Hepatocarcinoma Cells. PLoS One 2013;8. https://doi.org/10.1371/journal.pone.0076540.

[66] De Antonellis P, Carotenuto M, Vandenbussche J, De Vita G, Ferrucci V, Medaglia C, et al. Early targets of miR-34a in neuroblastoma. Mol Cell Proteomics 2014;13:2114–31. https://doi.org/10.1074/mcp.M113.035808.

[67] Chua CEL, Tang BL. miR-34a in Neurophysiology and Neuropathology. J Mol Neurosci 2019;67:235–46. https://doi.org/10.1007/s12031-018-1231-y.

[68] van den Berg MMJ, Krauskopf J, Ramaekers JG, Kleinjans JCS, Prickaerts J, Briedé JJ. Circulating microRNAs as potential biomarkers for psychiatric and neurodegenerative disorders. Prog Neurobiol 2020;185. https://doi.org/10.1016/j.pneurobio.2019.101732.

[69] Kim YJ, Kim SH, Park Y, Park J, Lee JH, Kim BC, et al. miR-16-5p is upregulated by amyloid β deposition in Alzheimer's disease models and induces neuronal cell apoptosis through direct targeting and suppression of BCL-2. Exp Gerontol 2020;136. https://doi.org/10.1016/j.exger.2020.110954.

[70] Everaert C, Luypaert M, Maag JLV, Cheng QX, DInger ME, Hellemans J, et al. Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. Sci Rep 2017;7. https://doi.org/10.1038/s41598-017-01617-3.

[71] Song L, Pan S, Zhang Z, Jia L, Chen WH, Zhao XM. STAB: A spatio-temporal cell atlas of the human brain. Nucleic Acids Res 2021;49:D1029–37. https://doi.org/10.1093/nar/gkaa762.

[72] Williams SE, Noel M, Lehoux S, Cetinbas M, Xavier RJ, Sadreyev RI, et al. Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. Nat Commun 2022;13. https://doi.org/10.1038/s41467-021-27781-9.

Chapter 2: Unique N-glycosylation signatures in Aβ oligomer- and lipopolysaccharide-activated human iPSC-derived microglia

Xinyu Tang[1]*, Ryan Lee Schindler[2]*, Jacopo Di Lucente[3], Armin Oloumi[2], Jennyfer Tena[2], Danielle Harvey[4], Carlito B. Lebrilla[2], Angela M. Zivkovic[1]**, Lee-Way Jin[3]**, Izumi Maezawa[3]**

[1]Department of Nutrition, University of California, Davis, California, USA.

[2]Department of Chemistry, University of California, Davis, California, USA.

[3]Department of Pathology and Laboratory Medicine and M.I.N.D. Institute, University of California Davis Medical Center, Sacramento, California, USA.

[4]Department of Public Health Sciences, University of California-Davis, Davis, CA, USA.

*Authors have equal contribution to the manuscript.

Address Correspondence to: Izumi Maezawa, Ph.D., Department of Pathology and Laboratory Medicine, University of California Davis Medical Center, Sacramento, CA 95817, USA, E-mail: imaezawa@ucdavis.edu, Lee-Way Jin, MD., PhD., Department of Pathology and Laboratory Medicine, University of California Davis Medical Center, Sacramento, CA 95817, USA, E-mail: lwjin@ucdavis.edu, Angela M. Zivkovic, PhD., Department of Nutrition, University of California, Davis, CA 95618, USA, E-mail: amzivkovic@ucdavis.edu.

ABSTRACT

N-glycosylation plays an important role in immune function, influencing the function of immune cells. Microglia are the immune cells in the central nervous system (CNS) and become pro-inflammatory/activated in Alzheimer's disease (AD). Cell surface glycosylation plays an important role in immune cells, however, the N-glycosylation and glycosphingolipid (GSL) signatures of activated microglia are poorly understood. Here, we study comprehensive combined transcriptomic and glycomic profiles using human induced pluripotent stem cells-derived microglia (hiMG). Distinct changes in N-glycosylation patterns in Aβ oligomer (AβO) and LPS -treated hiMG were observed. In AβO treated cells, the relative abundance of bisecting N-acetylglucosamine (GlcNAc) N-glycans decreased, corresponding with a downregulation of MGAT3, the gene responsible for bisecting GlcNAc N-glycan formation. The sialylation of N-glycans increased in response to AβO, accompanied by an upregulation of genes involved in N-glycan sialylation (ST3GAL2, 4, and 6). Moreover, we found that the N-glycosylation signature of LPS-induced hiMG differed from that of AβO-induced hiMG. LPS-induced hiMG exhibited a decreased abundance of complex-type N-glycans, aligned with downregulation of mannosidase genes (MAN1A1, MAN2A2, MAN1C1). Fucosylation increased in LPS-induced hiMG, aligned with upregulated fucosyltransferase 4 (FUT4) and downregulated alpha-L-fucosidase 1 (FUCA1) gene expression. However, the GSL profile did not exhibit significant changes in response to AβO or LPS activation. AβO- and LPS- specific glycosylation changes could contribute to impaired microglia function, highlighting glycosylation pathways as potential therapeutic targets for AD.

Keywords: Microglia, amyloid, lipopolysaccharide, glycosylation, sialylation, fucosylation

INTRODUCTION

Neuroinflammation plays significant roles in the initiation and progression of several central nervous system (CNS) disorders, including Alzheimer's disease (AD). The complexity of neuroinflammation reflects how microglia respond to internal and external changes to orchestrate inflammatory responses. Microglia are the innate immune cells of the CNS that survey and respond to various environmental challenges with multiple actions, resulting in phagocytosis of debris and release of cytokines and chemokines [1]. Prior to the development of specific immunohistochemical agents, a widely used classic histochemical method to detect microglia in brain sections was to employ lectins that recognize glycoconjugates [2], implying the significance of glycosylation (i.e., post-translational addition of glycans to protein backbones) in microglial function. Recently, the study of changes in glycosylation upon neuroinflammation is gaining increased interest [3]; however, how microglia respond to stimuli with glycosylation changes remains understudied.

AD pathology has highlighted the importance of microglia given their role in orchestrating neuroinflammation and phagocytosis. In AD brains, microglia lose their homeostatic molecular signature and show profound functional impairments, such as increased production of pro-inflammatory cytokines, elevated reactive oxygen species, impaired phagocytosis, and increased inflammasome formation [4]. Recently, the pivotal roles of microglia-orchestrated neuroinflammation in AD have been established [5–8]. A generally accepted hypothesis states that activated microglia become functionally impaired and release cytotoxic substances and pro-inflammatory cytokines that cause neuronal damage and aggravate AD pathology [8]. While multiple factors may cause microglial activation in AD, early studies have established that different species of Aβ aggregates are potent stimulants of microglia.

Among them, we and others found that the small soluble Aβ oligomer (AβO) assembled from Aβ42 peptide provides far stronger stimulation to induce microglial activation [9,10]. Aβ aggregates are recognized by a range of microglial pattern recognition receptors to induce mainly pro-inflammatory responses that could mediate Aβ-induced neurotoxicity, impair phagocytic function, and prime microglia to enhance their sensitivity and reactivity to inflammatory stimuli [11]. Understanding mechanisms of such activations could reveal microglial therapeutic targets.

In the present study, we characterized the global glycosylation changes of human microglia following AβO stimulation, using stimulation with lipopolysaccharide (LPS), a widely used microglia activation approach, for comparison. We employed human induced pluripotent stem cell (iPSC)-derived microglia (hiMG) as a major model rather than the widely used primary cultures from neonatal rodents based on the following rationales. First, it was recognized that rodents are generally a better model for neuronal pathology than they are for microglial pathology [12], and that for optimal translational validity, human microglia are recommended to be used to identify human-relevant molecular pathways and therapeutic targets [13]. Second, in contrast to the evolutionarily better conserved genetic codes and protein networks across species, there is no evidence for a universal "glycan code" akin to the genetic code or protein motif [14,15]. Rather, glycans vary immensely in structure and expression both within and between evolutionary lineages, and our knowledge about the glycan structural diversity between species remains limited [15]. Therefore, we consider that models of human cells would be suitable for our initial investigation on microglia glycosylation changes. Our profiling of hiMG identified the specific glycan types and glycosylation genes altered in AβO- and LPS-activated human microglia that may be involved in the regulation of their respective activation pathways.

METHODS

Human iPSC culture and microglia differentiation

Human iPSCs were obtained from ALSTEM.INC. (Richmond, CA). The line used in this study was Human iPS Cell Line 26 (Episomal, CD34+, ApoE3). Cells were plated onto Matrigel (Fisher) coated plates and cultured with mTeSR plus (Stemcell Technology). For microglia differentiation, we followed a previously described protocol [16,17]. Briefly, 2 x 106 iPSCs were plated onto Aggrewell 800 plates (Stemcell Technology) to form embryoid bodies (EBs) in mTeSR1 supplemented with Bone Morphogenetic Protein 4 (BMP4, 50 ng/ml)/ Vascular Endothelial Cell Growth Factor (VEGF, 50 ng/ml)/ Stem Cell Factor (SCF, 20 ng/ml) and culture for four days with daily medium change. On the fifth day, EBs were plated onto gelatin coated 6-well plates with 20 EBs per well in X-VIVO15 (Lonza) supplemented with M-CSF (100ng/ml), IL-3 (25ng/ml), Glutamax (2mM), Penicillin/streptomycin (100U/ml and 100ug/ml) and β-mercaptoethanol (55uM), and the medium was changed weekly. After 3-4 weeks, floating primitive macrophage precursor (PMP) were collected and plated onto 12-well plates (5 x 10^4 cells/well), 6-well plates (3 x 10^5 cells/well), or 100 mm dishes (1.5 x 10^6 cells) and differentiated in microglia differentiation medium (Advanced DMEM/F12 supplemented with IL-34 (100ng/ml), GM-CSF (10ng/ml), N2 supplement (1x), Glutamax (2mM), Penicillin/streptomycin (100U/ml and 100ug/ml) and β-mercaptoethanol (55uM)) for two weeks. The cultures were routinely tested for mycoplasma.

For AβO and LPS stimulation, cells were treated with AβO (3 μM) and LPS (100 ng/ml) for 24hrs. Each group, including the control group, has 6 replicates.

AβO and LPS preparation and treatment

AβO composed of Aβ1-42 peptide was prepared following a standard procedure [18] with a modification that the HFIP treated Aβ1-42 peptide (Bachem) was dissolved in DMSO and then diluted with Advanced DMEM/F12 culture medium instead of the F12 medium originally described, followed by incubation at 4°C for 24 hr and 10 min centrifugation at 10,000 x rpm at 4°C. This preparation of AβO has been extensively characterized in our laboratory [10]. Briefly, to ensure consistency of quality, a random sample from each batch was chosen and imaged using electron microscopy and atomic force microscopy to characterize the size and shape of the aggregates. The biological activities of each batch were confirmed by determining for AβO the neurotoxic activity, synaptic binding activity, and ability to rapidly induce exocytosis of MTT formazan, as described previously [10].

RNA isolation, processing, and sequencing

Total RNA from cultured cells was extracted using RNeasy Plus Mini Kit (Qiagen). RNA quality evaluation (yield, purity and integrity), cDNA library construction and Illumina sequencing were performed by NovoGene using 2 × 150 bp paired-end run (Sacramento, CA, USA). A total of 24 samples, comprising four groups (AβO vs Control and LPS vs Control) with six replicates per group, were sequenced with the output of 53.3±8.2 M reads (sequencing depth of 2.7±0.4) per sample.

Mapping and differential expression analysis

Adapter sequences in the raw FASTQ files were trimmed using Trimmomatic (v0.36). The remaining reads were mapped to the human reference hg38 using HISAT2 [19]. The mapping rate was around 97%. The gene counts were summarized by featureCounts [20], achieving a successful alignment of 70%-80% of the reads per library. Differential gene

expression was analyzed using edgeR package [21] in R version 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). Low-expressed genes were dropped using the *filterByExpr()* function. The filter step keeps genes that have count-per-million (CPM) above 10 in 70% of samples. The library sizes were recalculated after filtering. Sample- specific effects were removed by normalizing library sizes with the trimmed mean of M-values (TMM) method using the *calcNormFactors()* function. A negative binomial model with the quasi-likelihood (QL) F-test was applied to perform the differential expression analysis among diagnosis groups. First, gene counts with normalized library sizes were fitted to a negative binomial generalized linear model. Then, the dispersion was estimated using the *estimateDisp()* function. The QL dispersion estimation was calculated using the *glmQLFit()* function, followed by the *glmQLFTest()* function that conducted the quasi-likelihood (QL) F-test. Finally, we used the *topTags()* function to output significantly differentially expressed genes. P-values were adjusted for multiple hypothesis testing using Benjamini & Hochberg (BH) with a threshold of adjusted p-value $\leqslant 0.05$.

Cell Membrane Enrichment

Samples were prepared for glycomic analysis based upon previously published methodologies [22]. Briefly, harvested cells were centrifuged at 2,000 RCF for ten minutes after which the cell media was exchanged with 1.2 mL of 0.25 M sucrose (MilliporeSigma, Cat# S7903), 20mM HEPES (Thermo Fisher Scientific, 1 M, Cat# 15630080) adjusted to pH 7.4 with KOH (MilliporeSigma, Cat# P5958) and a 1:100 protease inhibitor cocktail (Calbio-chem, Cat# 539137) in water and pipette mixed gently. Cells were then lysed by sonication where a total of 60 joules of energy were applied to each sample over a one-minute period. The homogenate was centrifuged at 2,000 RCF for 10 minutes at 4°C and then the supernatant was transferred to 1.5 mL ultracentrifuge tubes and the nucleus precipitate was discarded. The cell membrane was

precipitated by ultracentrifugation at 42,000 RCF for 45 minutes at 4°C after which the cytoplasmic supernatant was discarded taking care not to disturb the membrane pellet. Samples were resuspended in 0.5 mL of 0.2 M Na2CO3 (MilliporeSigma, Cat# P5958) in water and ultracentrifuged again after which the supernatant, containing membrane-associated proteins, was discarded. A final ultracentrifugation step using the same conditions with 0.5 mL of water was done to wash and desalt the samples, discarding the supernatant after.

N-Glycan Sample Preparation

The enriched membrane fraction was then resuspended in 200 µL of 100 mM NH4HCO3 (MilliporeSigma, Cat# 09830), 5 mM Dithiothreitol (Promega, Cat# V3151) in water to assist in protein denaturing. After gentle homogenization by pipette mixing, samples were placed in a 100°C water bath for two minutes with 10 second heating and cooling intervals to further denature the membrane proteins. After samples cooled to room temperature, 2 µL of PNGase F (New England Biolabs, Cat #P0705L) was added to enzymatically hydrolyze the N-glycans from their protein substrates. Enzymatic cleavage and conversion of the resulting reducing end amides to hydroxyls occurred at 37°C for 18 hours. Sample were brought to a total volume of 600 µL with water and then ultracentrifuged using the previously described parameters. The supernatant containing the released N-glycans was transferred to a new vial and the remaining membrane pellet was stored at -20°C until ready for glycosphingolipid processing. N-glycan samples were cleaned up using a porous graphitic carbon solid phase extraction plate (PGC-SPE, Glygen, Cat# FNSCAR800). SPE wells were first washed with elution solvent consisting of 80% acetonitrile, 0.1% trifluoracetic acid in water (v/v) and then conditioned in wash solvent consisting of 40% acetonitrile, 0.05% trifluoracetic acid in water (v/v). Samples were loaded onto the solid phase and washed three times with 600 µL of washing solvent and all liquids were discarded. N-

glycans were then eluted with 600 µL of elution solvent into a new collection plate. For each SPE step, the PGC plate was centrifuged at 150 RCF to quickly ensure solvent flow-through. Enriched N-glycan samples were concentrated by drying to completion in vacuum centrifugation and reconstituted in 30 µL of water prior to analysis.

N-Glycan Analysis Methodology

Released N-glycan analysis was conducted with an Agilent 1200 series ChipCube system operating in positive ion mode and a 6520 Q-TOF mass spectrometer. Chromatographic separations were carried out utilizing a PGC-Chip II, 40 nL enrichment column, 75 x 43 mm analytical column, 5 µm (Agilent Technologies, G4240-62003). Mobile phase A (MPA) composition was 3% acetonitrile and 0.1% formic acid in water (v/v). Mobile phase B (MPB) composition was 90% acetonitrile, and 1% formic acid in water (v/v). Samples were cooled with a 4°C sample tray and 5µL was injected by the autosampler, loaded onto the enrichment column, and washed for 10 minutes with 100% MPA at 3µL/min. Next, the chip valve switched to analysis mode moving the enrichment column in line and the analytical gradient began with a flow rate of 0.33 µL/min. 0% MPB was held from 0 to 2 minutes, at 20 minutes MPB reached 16%, at 40 minutes MPB reached 72%, and at 42 minutes MPB reached 100%. The 100% MPB wash was held until 52 minutes and decreased back to 0% MPB at 54 minutes, equilibration of starting composition was held until 65 minutes. Source parameters used nitrogen as the drying gas with a temperature of 325°C flowing at 5L/min and a capillary voltage of 1800V. Auto MS/MS acquisition mode allowed 4 precursors per cycle with a threshold of 2000 counts and quadrupole filtering with a narrow isolation width for the 600-2000 m/z range. Collision-induced disassociation with N2 gas used a precursor-dependent formula y=1.8*(m/z)/100-3.6 to determine individual collision energies with a range of 100-2000 m/z. Continual mass correction

with a reference calibration standard of 1221.990637 m/z was used (Agilent Technologies, HP-1221). N-glycan compound verification used MassHunter Qual B08.00 software with the 'Find by Molecular Feature' algorithm and an in-house database containing N-glycan monoisotopic masses and descriptive information.

Sphingolipid Sample Preparation

The membrane pellet from the initial steps was removed from -20°C storage and allowed to acclimate to room temperature. Samples were resuspended in 500 µL of freshly prepared sphingolipid extraction solvent consisting of water/methanol/chloroform (3:8:4, v/v%) and sonicated for 15 minutes to assist in solvation. Proteins were precipitated by centrifugation at 8,800 RCF for 5 minutes and the supernatant was transferred to a new 1.5 mL Eppendorf vial. A bilayer phase separation was induced by the addition of 100 µL of 0.1 M KCl (MilliporeSigma, Cat#104936). The top aqueous-rich layer containing the sphingolipids of interest was transferred to a new vial and dried by vacuum centrifugation. Sphingolipid samples were cleaned up using a C-8 solid phase extraction plate (Glygen, Cat# FNSC08.800). SPE wells were first washed with elution solvent consisting of methanol/isopropyl alcohol (1:1, v/v%) and then conditioned in wash solvent consisting of methanol/water (1:1, v/v%). Samples were reconstituted with wash solvent, loaded onto the solid phase by gravity and the flow-through was reloaded. Sample wells were washed three times with 600 µL of washing solvent and all liquids were discarded. Sphingolipids were then eluted with 400 µL of elution solvent into a new collection plate. For each SPE step, the C-8 plate was allowed to gravity load to ensure analyte binding. Enriched sphingolipid samples were concentrated by drying to completion in vacuum centrifugation and reconstituted in 20 µL of water/methanol (1:1, v/v%) prior to analysis.

Sphingolipid Analysis Methodology

Sphingolipid analysis was conducted with an Agilent 1200 series nanoflow HPLC equipped with a 10pt/2ps μ-valve(G1316-68709), nanoflow ESI operating in positive ion mode (G1992A), and a 6530 Q-TOF mass spectrometer. Chromatographic separations were carried out utilizing a Zorbax SB-300 C-8, 0.075 x 50 mm, 3.5 μm column (Agilent Technologies, Cat# 5065-9923). Four mobile phase compositions were used in this analysis with two low-flow pumps. The first pump (pump 1) operated at 0.7 μL/min and was used to load samples directly and wash the analytical column. The loading mobile phase (MPL) composition was 20 mM ammonium acetate in 56% methanol, 14% isopropyl alcohol, 0.1% glacial acetic acid and water (v/v). The washing mobile phase (MPW) composition was 50% methanol and 50% isopropyl alcohol (v/v). The second pump (pump 2) operated at 0.4 μL/min and was used to carry out the analytical gradient. Mobile phase A (MPA) had a composition of 20 mM ammonium acetate in 0.1% glacial acetic acid and water (v/v). Mobile phase B (MPB) had a composition of 20 mM ammonium acetate in %0.1% glacial acetic acid, 20% isopropyl alcohol, and 80% methanol (v/v). Samples were cooled with a 4°C sample tray and 1μL was injected and loaded directly onto the analytical column for 4 minutes with 100% MPL. At 4 minutes the μ-valve switched positions moving the analytical column into pump 2's flow path and the gradient began. 70% MPB was held from 4 to 6 minutes, at 35 minutes MPB reached 93%, and at 36 minutes MPB reached 100%. The 100% MPB was held until 39 minutes and decreased back to 70% at 40 minutes, equilibration of the starting composition was held until 47 minutes. Source parameters used nitrogen as the drying gas with a temperature of 300°C flowing at 5L/min and a capillary voltage of 1000V. Auto MS/MS acquisition mode allowed 5 precursors per cycle with a threshold of 600 counts and quadrupole filtering with a narrow isolation width for the 600-2000 m/z range. Collision-induced disassociation with N2 gas used a precursor-dependent formula

y=1.2*(m/z)/100+12 to determine individual collision energies with a range of 100-2000 m/z. Continual mass correction with a reference calibration standard of 1221.990637 m/z was used (Agilent Technologies, HP-1221). Sphingolipid compound verification used MassHunter Qual B08.00 software with the 'Find by Molecular Feature' algorithm and an in-house database containing sphingolipid monoisotopic masses and descriptive information.

Statistical analysis

Single glycans with missing values for more than 80% of samples were excluded from further analysis (**Supplementary Figure 2-S1**). The fractions of glycan subtypes (as shown in **Figure 2-S1**), and sialylated and fucosylated N-glycans were calculated as the sum of relative abundances of single glycans. Differential abundance analysis was performed as linear regression models using the lmerTest package in R [23] as shown here: glycan abundance ~treatment. ANOVA tests and contrasts between treatments and control were evaluated. The AβO and LPS groups were compared to the control group using Dunnett's test. The differences of the three treatment groups on the PCA plot were evaluated using the MANOVA test followed by post-hoc linear discriminant analysis (LDA).

RESULTS

Both AβO and LPS induced significant changes in the N-glycan profiles of hiMG

We analyzed the N-glycan profiles of control (mock treatment), AβO-activated, and LPS-activated hiMG using LC-MS/MS technology, with six replicates per group. We treated hiMG for 24 hr with 3 μM AβO and 100 ng/ml LPS, optimal concentrations to induce pro-inflammatory cytokines IL-1β, TNF-α, and IL-6 (Jin et al., 2023). **Figure 2-1** illustrates the N-linked glycosylation types. In all groups, the complex-type N-glycans were the most abundant N-glycans, making up ~35-40% of total N-glycans (**Figure 2-2A**). This is in keeping with the general composition of brain glycans [24]. Hybrid glycans were the least abundant in hiMG, making up less than 10% of total N-glycans (**Figure 2-2A**). Complex-type N-glycans can be further classified as bi-, tri, and tetra-antennary, as well as bisecting GlcNAc, which contains an additional GlcNAc bound to the core (**Figure 2-1**). In hiMG, the tetra-antennary structure constituted the predominant subtype of complex-type N-glycans, followed by bi-antennary, bisecting GlcNAc, and tri-antennary structures (**Figure 2-2B**). N-Glycans can be decorated by adding fucose, sialic acid, or both to the terminal end of glycans. Fucose can also be added to the core structure. These modifications are important for protein functions such as protein stability, cell adhesion, signal transduction, immunological responses, and cell-to-cell interactions [25,26]. We quantified such modifications and found that in hiMG, sialofucosylated N-glycans were the most abundant, making up ~50% of glycans. About 30% of N-glycans were not decorated by fucose or sialic acid residues. The rest of N-glycans were only decorated by either fucose or sialic acid residues (**Figure 2-2C**). Considering both subtypes and decorations, sialofucosylated complex N-glycans and undecorated high-mannose N-glycans are the two most abundant structures in hiMG (C-FS and HM in **Figure 2-2D**). AβO and LPS treatments had a significant

impact on the N-glycan profile of hiMG as shown by the increased dispersion of the AβO-activated hiMG compared to the control cluster and the separation of the LPS-activated hiMG clusters from the control cluster on the PCA plot (MANOVA p-value = 0.0014). The distance between the LPS cluster and the control cluster was greater than the distance between the AβO cluster and the control cluster, indicating that LPS might induce a higher degree of glycosylation changes than AβO (**Figure 2-2E**, **Supplementary Figure 2-S2A**).

Increased sialylation and decreased bisecting GlcNAc N-glycans in AβO-activated hiMG

According to the N-glycome, AβO, but not LPS, treatment reduced the percent abundance of bisecting GlcNAc N-glycans in hiMG (**Figure 2-2B**). Consistent with this glycomic change, differential expression (DE) analysis of RNA-seq data revealed that following AβO treatment, beta-1,4-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase 3 (MGAT3) was the most downregulated gene across all genes involved in the core N-glycan formation process (**Figure 2-3A**). MGAT3 encodes N-acetylglucosaminyltransferase-III (GnT-III), which is uniquely responsible for transferring a GlcNAc residue to the mannose of the trimannosyl core of N-glycans to produce a bisecting GlcNAc N-glycan [27] (**Figure 2-3D**). Thus, the downregulation of the MGAT3 gene could be the transcriptional basis of the reduction of bisecting GlcNAc N-glycans observed in AβO treated hiMG cells.

AβO treatment also increased the percent abundance of sialylated N-glycans (**Figure 2-2C**), in particular, sialylated complex N-glycans (C-S in **Figure 2-2D**). We therefore further explored the changes in the number of sialic acid residues on N-glycans and found that the increased sialylation was largely attributed to an increased percent abundance of di-sialylated N-glycans (**Figure 2-3B**). DE analysis of RNA-seq data showed that among the genes encoding sialyltransferases and sialidases, ST3GAL2, ST3GAL4, and ST3GAL6 were significantly

upregulated by AβO treatment (**Figure 2-3C**). Given that ST3GALs encode for enzymes responsible for transferring sialic acid to the galactose residues with α-2-3 linkage (**Figure 2-3D**), the results suggest that α-2,3-linked sialic acids possibly contributed to the increased di-sialylation observed in the AβO-activated hiMG.

DE analysis also revealed increased levels of FUT8 transcript encoding α1-6 fucosyltransferase (fucosyltransferase 8), which is uniquely responsible for core fucosylation, in AβO-treated hiMG (**Figure 2-3A**). This result corroborates our published data reporting that AβO enhances FUT8-catalyzed core fucosylation, which is a signaling pathway required for AβO-induced microglia activation [16].

Decreased complex N-glycans and increased fucosylation in LPS-activated hiMG

In contrast to AβO-activated hiMG, LPS-activated hiMG exhibited a reduced percent abundance of complex-type N-glycans, specifically the tetra-antennary complex N-glycans (**Figure 2-2A, B**). Complex N-glycans are derived from high mannose N-glycans. In the endoplasmic reticulum (ER) and cis-Golgi, α-mannosidase I (MAN1B1), α1-2 mannosidases IA, IB, and IC (encoded by MAN1A1, MAN1A2, and MAN1C1, respectively) trim mannose residues from high mannose N-glycans to yield Man5GlcNAc2, a key intermediate in the pathway to generate hybrid and complex N-glycans (**Figure 2-4E**). In the DE analysis of hiMG transcriptomic data, MAN1A1, MAN1C1, and MAN2A2 were significantly downregulated in LPS-treated hiMG (**Figure 2-4A**). The transcript levels of other key mannosidases, including MAN1B1, MAN1A2, and MAN2A1, were also decreased though not statistically significant (**Figure 2-4A**). The downregulation of these mannosidases may contribute to the reduced abundance of complex N-glycans.

The changes in LPS-activated hiMG also differed from AβO-activated hiMG. The extent of fucosylation (i.e., fucosylated-only N-glycans), especially complex- and hybrid-type N-glycans, was increased in LPS-activated hiMG but remained unchanged in AβO-activated hiMG (**Figure 2-2C-D**). This overall change was the sum of changes of individual glycan structures. In contrast, sialofucosylated N-glycans, the most abundant types, were significantly reduced overall in LPS-activated hiMG, especially sialofucosylated complex N-glycans (**Figure 2-2C-D**). Interestingly, among the sialofucosylated N-glycans those containing one fucose residue decreased while those containing two fucose residues increased (**Figure 2-4C**). Thus, the decrease in mono-fucosylated and concomitant increase in di-fucosylated complex-type N-glycans also containing sialic acid was the primary effect of LPS treatment.

Fucosyltransferases (FUTs), except FUT8, are responsible for transferring fucose to antenna GlcNAc residues (**Figure 2-4D**). According to the DE analysis, FUT4 was significantly upregulated in LPS-treated hiMG, which possibly contributed to increased fucosylation, especially increased di-fucosylated N-glycans. Fucose residues on N-glycans can be hydrolyzed by fucosidases (FUCAs). Alpha-L-Fucosidase 1 (FUCA1) specifically hydrolyzes the fucose residue at the core structure (**Figure 2-4E**). The transcript level of FUCA1 was significantly downregulated (**Figure 2-4D**), suggesting less hydrolysis of core fucosylated N-glycans. This, together with enhanced FUT4 action, could lead to increased percent abundance of fucosylated N-glycans.

The glycosphingolipid (GSL) profile of hiMG

GSLs are a subclass of lipids composed of a ceramide part and a mono- or oligosaccharide part. In the adult mammalian brain, the major GSLs are the gangliosides of GM1, GD1a, GD1b, and GT1b that have one to three sialic acid residues respectively [28,29].

Because sialic acid residues on a glycoprotein or glycolipid may modulate cell signaling mediated by lectins such as selectins and sialic acid-binding immunoglobulin-type lectins (Siglecs) and because we found significant changes in sialic acid levels following AβO and LPS treatment, we profiled the GSL composition as well. The most abundant GSL was ganglioside GM3, the precursor of other complex gangliosides (**Figure 2-5A**). The second most abundant species was sphingomyelin (SM), a ceramide with a phosphocholine headgroup, followed by neutral GSL G2. The remaining GSLs were less abundant with each accounting for less than 5%. Regarding sialylation, GSLs detected in hiMG were predominantly sialylated GSLs (**Figure 2-5B**). Interestingly, the GSL profile did not change significantly in response to either AβO or LPS activation. The PCA plot showed that the clusters of control, AβO, and LPS groups did not separate from each other (MANOVA p-value = 0.56, **Figure 2-5C**, **Supplementary Figure 2-S2B**). The genes involved in GSL biosynthesis also did not show significant changes in response to AβO activation (**Supplementary Figure 2-S3A**). In LPS-activated hiMG, even though the genes involved in GSL biosynthesis, i.e., UGCG, B4GALT5, and B4GALNT1, were significantly upregulated, no significant changes were observed when individual GSLs in either treatment were examined (**Figure 2-5A, Supplementary Figure 2-S3B**).

DISCUSSION

The composition and balance between various types of glycans play an important role in regulating cell functions [30]. Cellular N-glycans can be subclassified as complex-, high mannose-, and hybrid-type N-glycans. High-mannose N-glycans are the precursor of hybrid and complex-type N-glycans. Mannosidases hydrolyze mannose from the high-mannose structure, preparing it for complex- and hybrid-type N-glycan biosynthesis. High mannose glycans have been reported as the major N-glycan type in the undifferentiated human monocytic leukemia cell line THP-1, primary blood-derived CD14+ monocytes, and in Caco-2 cells [31–34]. After differentiation, the levels of high mannose glycans in these cells decreased, accompanied by an increase in complex-type structures. The latter became the predominant N-glycan type in differentiated cells. In line with the differentiated state of microglia, hiMG expresses predominantly complex-type N-glycans regardless of activation. On the other hand, our hiMG data (summarized below), when compared to published data from other cell types, show differences mainly in levels of glycan decorations such as mono- or di-sialylation, fucosylation, or sialofucosylation, suggesting their significance in regulating cell-specific functions.

The differential glycosylation changes stimulated by AβO and LPS are summarized in **Figure 2-6**. . Congruent evidence from both glycomic and transcriptomic analyses reveals the transcriptional basis of these changes and enhances the validity of our findings. We found that AβO induces increased sialylation catalyzed by ST3GALs and decreased bisecting GlcNAc N-glycans catalyzed by GnT-III. A third major glycosylation change, previously reported by us [16], is increased core fucosylation catalyzed by FUT8, which is corroborated in the current study by increased expression of FUT8 transcript. These three glycosylation changes are known to affect immune regulation [25,30,35], suggesting their significance in microglial function. They

have also been implicated in AD, as published data using brain or cerebrospinal fluid samples from participants with AD or mild cognitive impairment (MCI) highlighted altered levels of bisecting GlcNAc, sialylation, and core fucose in N-glycans [3,36–39]. Altered glycoforms of individual AD-relevant proteins were also reported; for example, the familial AD mutant of the Aβ-amyloid precursor protein (APP) was shown to have a higher content of bisecting GlcNAc and core fucose residues compared to wild-type APP [40]. However, exact glycosylation-regulated molecular mechanisms, especially those pertaining to neuroinflammation in AD are largely unknown. Furthermore, a widely recognized feature of glycobiology is that different cell types express distinct glycosyltransferase isozymes, and thus proteins may have different glycosylation patterns depending on their cell of origin and in response to functional needs [41]. Therefore, analyses of multicellular brain tissues or other cell types do not precisely reflect glycosylation alterations in microglia. In this context, the microglia- and stimulant-specific glycosylation changes we report here would provide the first significant insights into how glycosylation alterations may regulate microglial function.

Among the three specific glycosylation changes, sialylation is the most studied in AD due to the high level of interest in Siglec, in part based on the association of the gene of CD33, a Siglec receptor, to AD risk [42]. Neuronal membranes contain a high density of sialic acid residues on glycoproteins, which protect neurons from aberrant microglial phagocytosis [43] via interacting with Siglec receptors such as CD22 and CD33 on microglia [44–46]. However, how sialylation changes in microglia *per se* affect neuron-microglia interaction or microglial function is poorly understood. Previous studies using the mouse BV2 microglia cell line found that BV2 cells activated by LPS, fibrillar Aβ, and tau had increased sialidase activity that desialylated the BV2 cell surface, promoting phagocytosis of neurons mediated by CR3 [44,47]. Although we

found that oligomeric Aβ-activated hiMG showed an increase in total sialylation, the RNA-seq analysis suggests that the changes in sialylation may be specific to glycosidic linkages, with a decrease in α2-6 and α2-8 linked sialic acids and an increase in α2-3 linked sialic acids. In regard to CD33 ligand binding preference, evidence points to overlapping specificities for α2-3 and α2-6 linked sialic acids [48] where the binding affinity was increased with sulfation [49,50] and decreased with branched α1-3 fucosylation [51]. Of note, AβO treatment of hiMG was accompanied by reduced phagocytosis of pHrodo Green BioParticles and of Aβ itself as demonstrated previously [16], revealing the complexity and need for further investigation. The differences may be explained by different cell models (BV2 vs. hiMG), interspecies differences (mouse vs. human), different modes of microglia activation (fibrillary Aβ or LPS vs AβO), and different methods of analysis. We used human iPSC-derived microglia as the model due to its strong human relevance, while the BV2 model is a rapidly growing tumor cell line with well recognized limitations [52]. AβO is considered a primary pathological agent in early stages of AD preceding Aβ fibril formation [53], and the respective influences on microglia state by AβO and Aβ fibril may represent different neuroinflammation phases in AD.

Regarding increased core fucosylation, our recent study using the same hiMG model provides the first indication about its significance in microglia activation [16]. We further found that FUT8 expression was increased in both human AD brains and microglia isolated from 5xFAD mice, a model of AD-like cerebral amyloidosis. Moreover, FUT8 is a component of the p53 signaling cascade regulating microglial behavior; FUT8-catalyzed core fucosylation is required for AβO-activated microglial alterations, including induction of pro-inflammatory cytokines, activation of p38MAPK, and phagocytic deficits [16].

Bisecting GlcNAc, a β1,4-linked GlcNAc attached to the core β-mannose residue, is a structure specific to complex N-glycans. It is a prominent feature of human brain complex N-glycans, occurring at a frequency of 40% [54]. The presence of a bisecting GlcNAc on glycoproteins has many implications in biological functions, such as in immune tolerance [55], tumor metastasis, and brain development [56,57]. The relevance of bisecting GlcNAc to AD is suggested by its altered levels in AD brain samples together with altered expression of GnT-III mRNA. However, whether it is up- or down-regulated in AD remains inconclusive as conflicting results were reported [36–38]. Interestingly, APP and β-site APP cleaving enzyme-1- (BACE1), two key proteins required for Aβ production, contain glycosylation modifications with bisecting GlcNac. It was reported that Aβ (aggregation state not characterized) treatment enhanced GnT-III mRNA expression in Neuro2a mouse neuroblastoma cells. This was considered neuroprotective as GnT-III-transfected cells showed increased α-secretase activity and decreased production of Aβ40 and Aβ42 [37]. Apparently contradicting this notion, Kizuka et al. reported that modifications with bisecting GlcNac stabilized BACE1 to increase Aβ production and that GnT-III deficiency reduced Aβ-plaque formation in the brain by accelerating lysosomal degradation of BACE1 [58]. Again, the significance of altered bisecting GlcNac modifications in microglia is poorly understood. Our observation of decreased GnT-III mRNA expression and decreased bisecting GlcNac abundance in AβO-activated hiMG provides the first indication that bisecting GlcNac modifications play a role in regulating microglial function.

Compared to AβO, LPS clearly induced a distinct glycosylation-related transcriptomic and glycomic pattern with decreased abundance of complex N-glycans and increased abundance of fucosylation, consistent with their different signaling pathways. The combination of increased FUT4 and decreased FUCA1 expression may underlie the increased overall fucosylation we

found in LPS-activated hiMG. However, in contrast to AβO effects, we did not find increased core fucosylation following LPS treatment [16], suggesting that LPS mainly enhances terminal fucosylation. In addition, the fucosylation changes were differential among glycan structure subgroups with the overall increase in fucosylation consisting predominantly of increased fucosylated-only N-glycans and decreased mono-fucosylated complex N-glycans also containing sialic acid. The functional implications of these novel structural changes warrant further investigation. Our results also prompt a novel hypothesis that LPS downregulates mannosidases, which trim mannose residues from high mannose N-glycans to yield Man5GlcNAc2, a key intermediate in the pathway to generate hybrid and complex N-glycans, resulting in reduced levels of complex N-glycans. This hypothesis would need further investigation.

Little is known about GSL in microglia, despite reported changes in ganglioside profiles in neurodegenerative disorders [59] and the anti-inflammatory effects of GM1 and other gangliosides on microglia [60]. It was observed by dot blot analysis that microglia were characterized by abundant GM1 in mixed murine glia culture, and that treatment of mouse primary microglia with LPS resulted in decreased GM1 and GT1b levels [60,61]. However, these previously reported changes were not apparent in our GSL profiling of hiMG. We identified no GSL alterations in hiMG following AβO or LPS stimulation. The most abundant GSL was ganglioside GM3, the precursor of other complex gangliosides (**Figure 2-4A**). The second most abundant sphingolipid was SM followed by ganglioside G2. The remaining GSLs were less abundant with each accounting for less than 5%. When considering the number of sialic acids, GSLs detected in hiMG were predominantly sialylated GSLs (**Figure 2-4B**).

It's important to note some limitations in this study. While human iPSC is highly similar to microglia *in vivo* in terms of transcriptome and morphology, it still doesn't fully replicate

microglial responses *in vivo* [62]. In the brain, various cell types, particularly neurons, could interact with microglia in response to AβO, but our study doesn't capture these cell-cell interactions. Moreover, microglia can exhibit heterogeneity driven by factors such as brain regions and sex, which isn't reflected in our *in vitro* model [63,64]. Lastly, although we observed significant glycosylation changes in microglia in response to LPS, it's worth noting that LPS itself isn't commonly found in the brain.

In summary, using a human microglia culture model, we were able to identify differential AβO- and LPS-induced glycosylation changes that may impact the functional or cellular interaction outcomes of these two types of pro-inflammatory activation, and we did not identify significant GSL changes. Our data add to currently scant information about microglia- and stimulant-specific "glycosylation codes". They also help generate hypotheses that may lead to a better glycoproteomic and glycolipidomic understanding of microglial function, which is exponentially more complex than the glycomic landscape we report here.

DECLARATIONS

Availability of data and materials

The data that support the findings of this study are available in the methods and/or supplementary material of this article, and are also available from the corresponding author upon reasonable request. The RNA sequencing data will be submitted to the NCBI Gene Expression Omnibus repositories. N-glycomic and glycosphingolipid data are available upon request after acceptance.

Author contributions

Izumi Maezawa, Lee-Way Jin, Carlito Lebrilla, Angela Zivkovic, Danielle Harvey: designed the study; Xinyu Tang, Izumi Maezawa, Lee-Way Jin, Ryan Lee Schindler, Angela Zivkovic wrote the article; Lee-Way Jin, Jacopo di Lucente, Ryan Lee Schindler, Armin Oloumi, Jennyfer Tena, and Izumi Maezawa: performed the experiments; Xinyu Tang, Ryan Lee Schindler, and Danielle Harvey: RNA-seq data and glycomic data analysis; Carlito Lebrilla: provided reagents, technologies.

Competing interests

All authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by National Institute on Aging awards R01 AG062240 and RF1 AG071665.

TABLES AND FIGURES

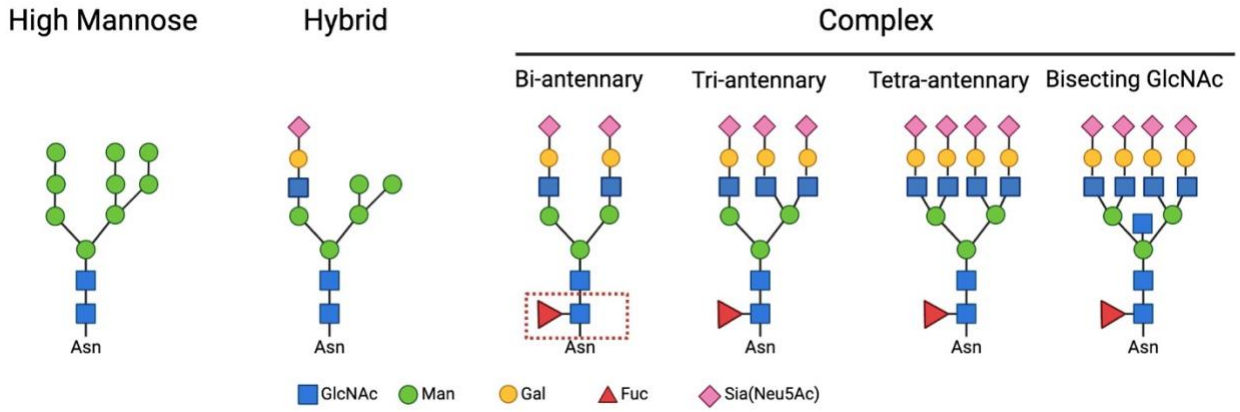


Figure 2-1 Outline of N-glycan classes. Examples from left to right represent high mannose, hybrid, and complex N-glycans. Complex N-glycans can be classified as bi-, tri-, tetra-antennary, and bisecting GlcNAc based on the number of starting GlcNAc residues on the mannose residues. The core fucosylation is marked with a red dash square. Recommended symbol nomenclature was used. GlcNAc: N-acetylglucosamine; Man: mannose; Gal: galactose; Fuc: fucose; Sia: sialic acid.

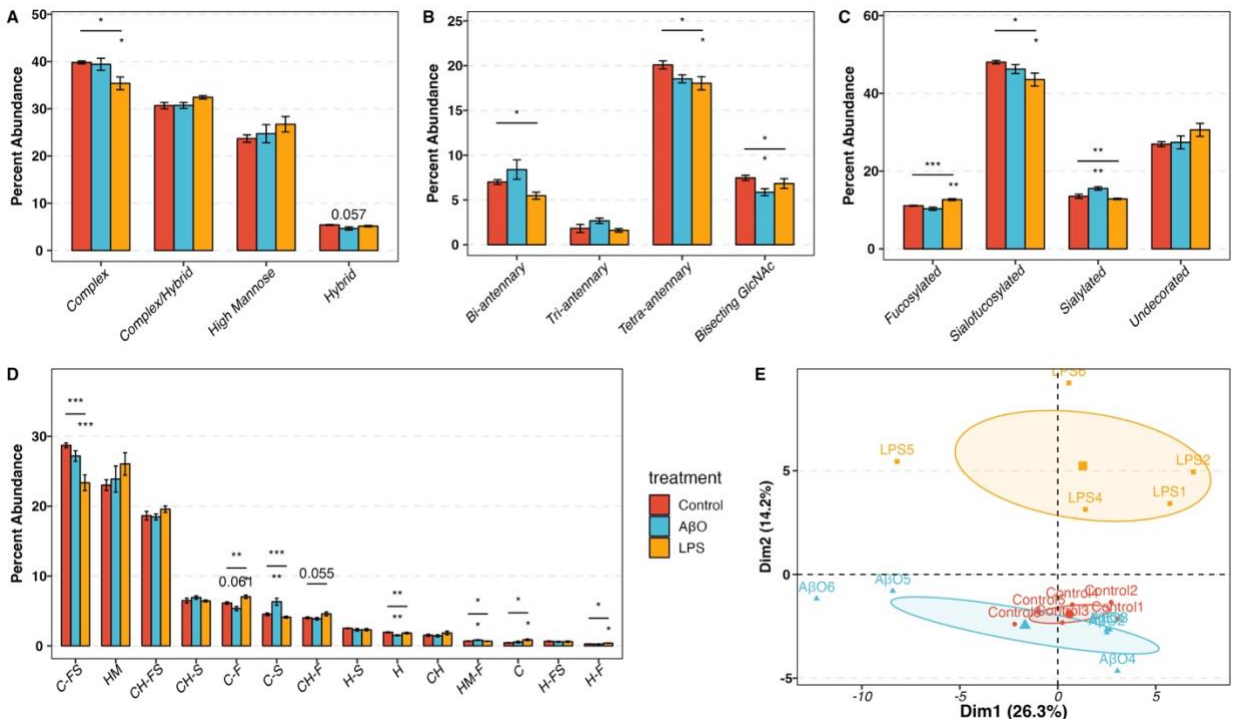


Figure 2-2 N-glycomic profile in control, AβO-, or LPS-activated hiMG. **A**) the percent abundance of complex, complex/hybrid, high mannose and hybrid N-glycan. **B**) The percent abundance of bi-, tri-, tetra-, and bisecting complex N-glycan. **C**) The percent abundance of each decoration, including fucosylated, sialofucosylated, sialylated, and undecorated N-glycan. **D**) The percent abundance of each decoration on each N-glycan type. Glycans are named as "glycan type-decoration". C, CH, H, HM represent Complex, Complex/Hybrid, Hybrid, and High Mannose, respectively. S, F, SF represent sialylated, fucosylated, and sialofucosylated, respectively. **E**) Principal Component Analysis (PCA) plot showing the clustering of control, AβO- and LPS-activated samples. MANOVA p-value =0.0014. Red, blue, and yellow correspond to control, AβO, and LPS groups. The asterisk (*) marks above the segments indicate the p-values obtained from the one-way ANOVA test. The asterisk marks below the segments indicated

the p-values from the Dunnett test for the corresponding treatment group. p values *<0.05, **<0.01, and ***<0.001.

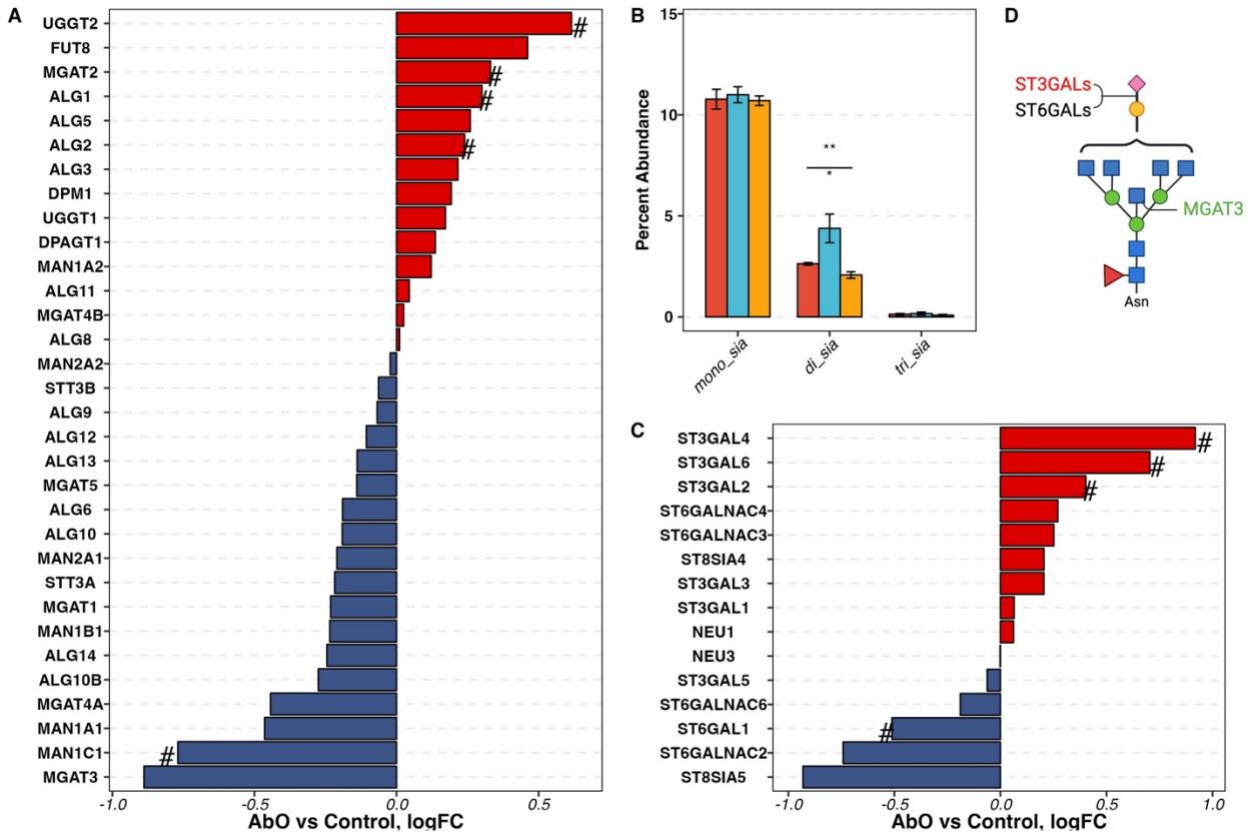


Figure 2-3 Sialylation changes in AβO-activated hiMG. **A**) The differential expression of genes involved in the biosynthesis of N-glycan core structure. **B**) The percent abundance of mono-, di-, and tri- sialic acid residues among sialylated only N-glycans. **C**) The differential expression of genes involved in sialylation. **D**) The schematics of the N-glycan structure and the corresponding enzymes. Red, blue, and yellow denote control, AβO, and LPS groups. For **A**) and **C**), the pound (#) denotes genes with unadjusted p-value < 0.05 but FDR >0.05. For B), red, blue, and yellow correspond to control, AβO, and LPS groups. The asterisk (*) marks above the segments indicate the p-values obtained from the one-way ANOVA test. The asterisk marks below the segments indicated the p-values from the Dunnett test for the corresponding treatment group. p values *<0.05, **<0.01, and ***<0.001.

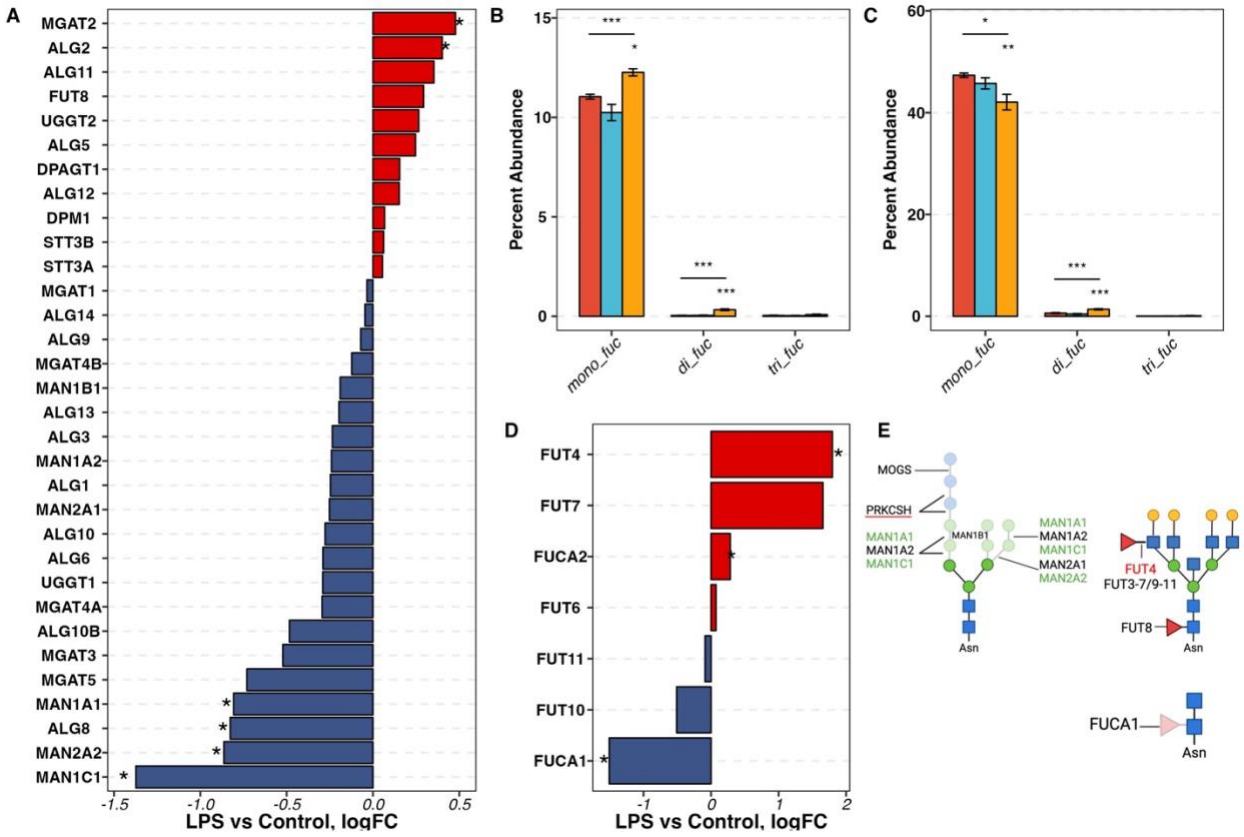


Figure 2-4 Fucosylation changes in LPS-activated hiMG. **A**) The differential expression of genes involved in the biosynthesis of N-glycan core structure. **B**) The percent abundance of mono-, di-, and tri-fucose residues among fucosylated only N-glycans. **C**) The percent abundance of mono-, di-, and tri-fucose residues among sialofucosylated N-glycans. **D**) The differential expression of genes involved in fucosylation. **E**) The schematics of complex N-glycan biosynthesis and fucosylation, including the enzymes involved. Red, blue, and yellow denote control, AβO, and LPS groups. For A) and D) The asterisk (*) denotes genes with FDR < 0.05. For B) and C), red, blue, and yellow correspond to control, AβO, and LPS groups. The asterisk (*) marks above the segments indicate the p-values obtained from the one-way ANOVA

test. The asterisk marks below the segments indicated the p-values from the Dunnett test for the corresponding treatment group. p values *<0.05, **<0.01, and ***<0.001.

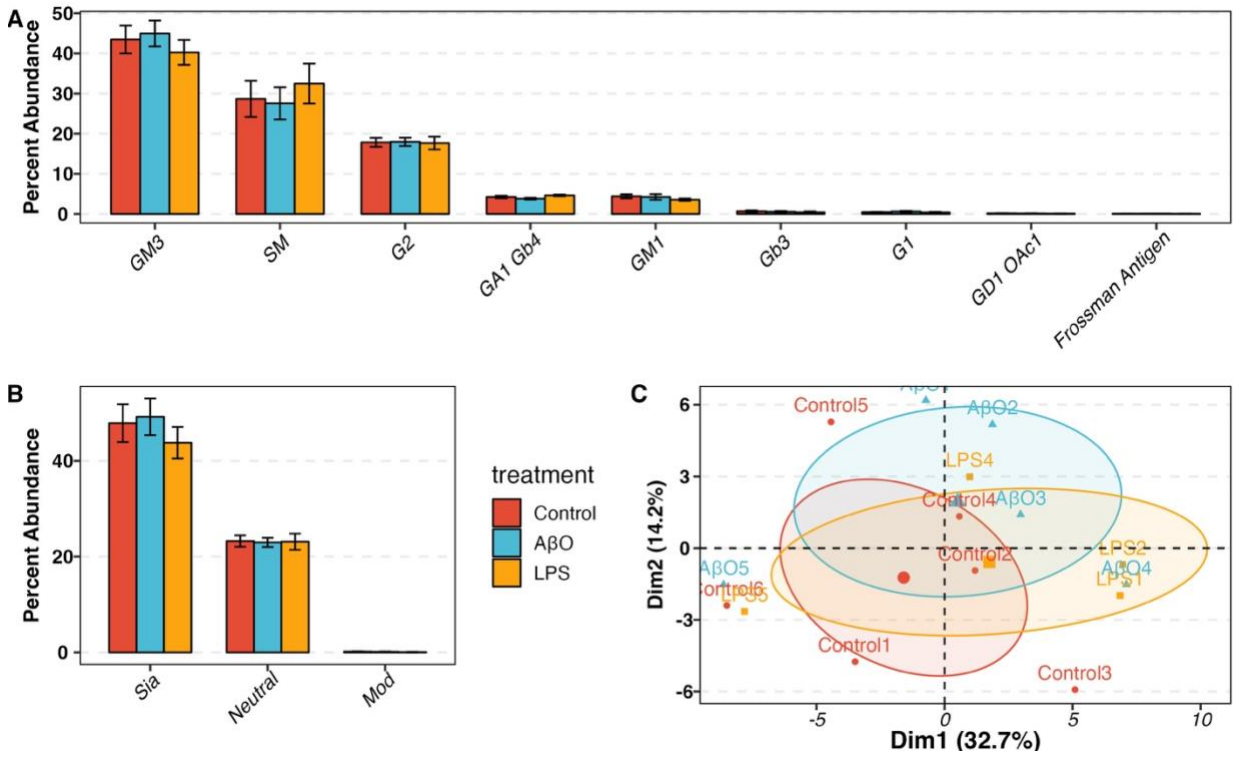


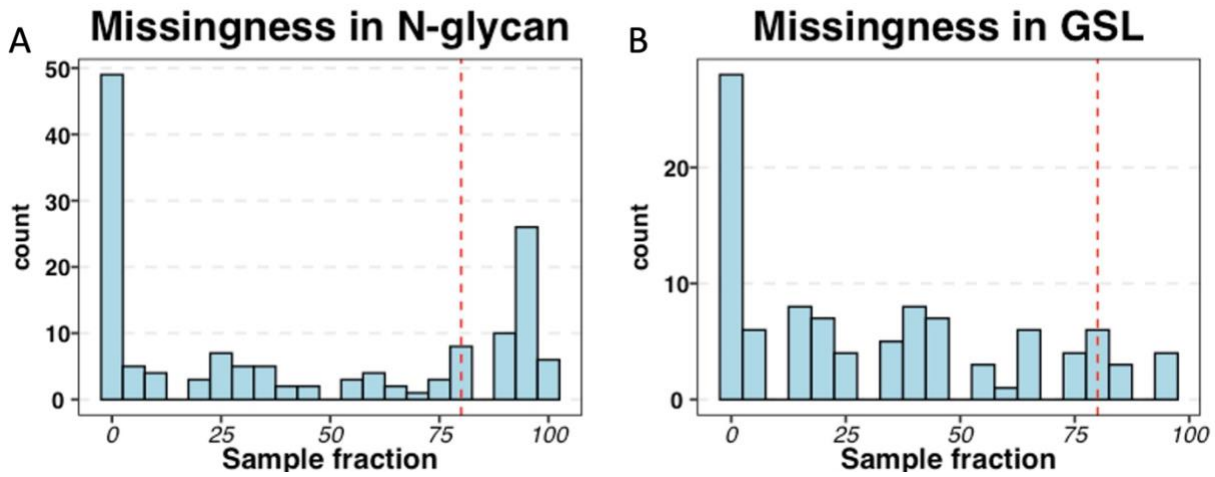
Figure 2-5 Glycosphingolipid (GSL) profile in control and AβO-, or LPS-activated hiMG. **A**) The percent abundance of GSL subtypes. **B**) The percent abundance of decorated GSLs. **C**) Principal Component Analysis (PCA) plot showing the clustering of control, AβO- and LPS-activated samples. MANOVA p-value =0.56. Red, blue, and yellow correspond to control, AβO, and LPS groups.

Figure 6.

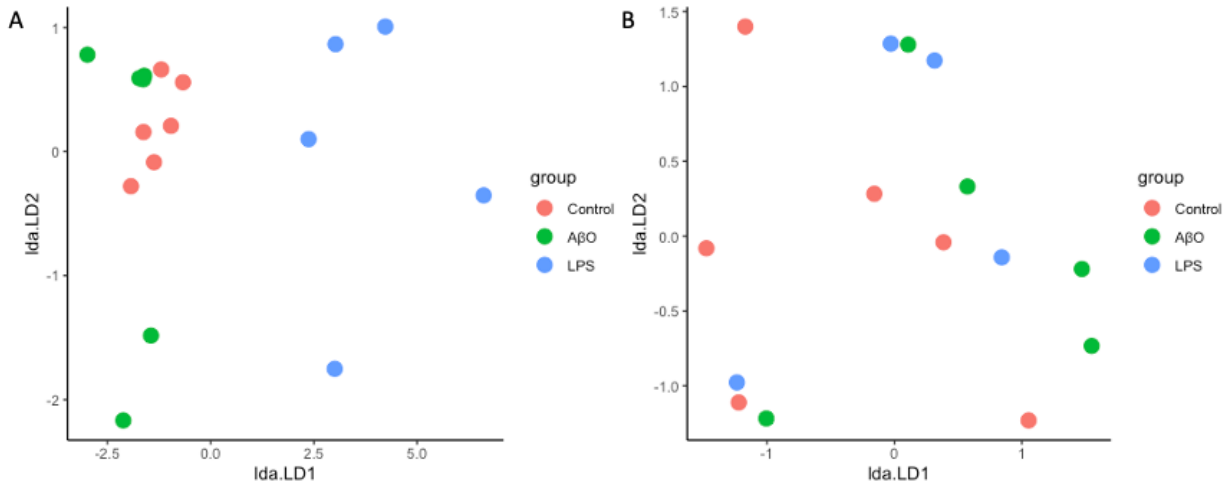


	Glycan type/Decoration	Example	Associated genes
AβO	Complex/Bisecting GlcNAc ↓	C-F_7710 	MGAT3 ↓
	Complex/Di-sialylated ↑	C-S_5402 	ST3GAL2 ↑ ST3GAL4 ↑ ST3GAL6 ↑
	Complex/Core fucosylated ↑*	C-FS_6513 	FUT8 ↑
LPS	Complex (tetra-antennary) ↓	C-FS_7611 	MAN1A1 ↓ MAN1C1 ↓ MAN2A2 ↓
	Complex or hybrid/Fucosylated ↑ - Mono-fucosylated ↑	C-F_7610 	FUT4 ↑ FUCA1 ↓
	- Di-fucosylated ↑	C-F_7620 	
	Complex/Sialofucosylated ↓ - Mono-fucosylated ↓	C-FS_7611 	
	- Di-fucosylated ↑	C-FS_7621 	

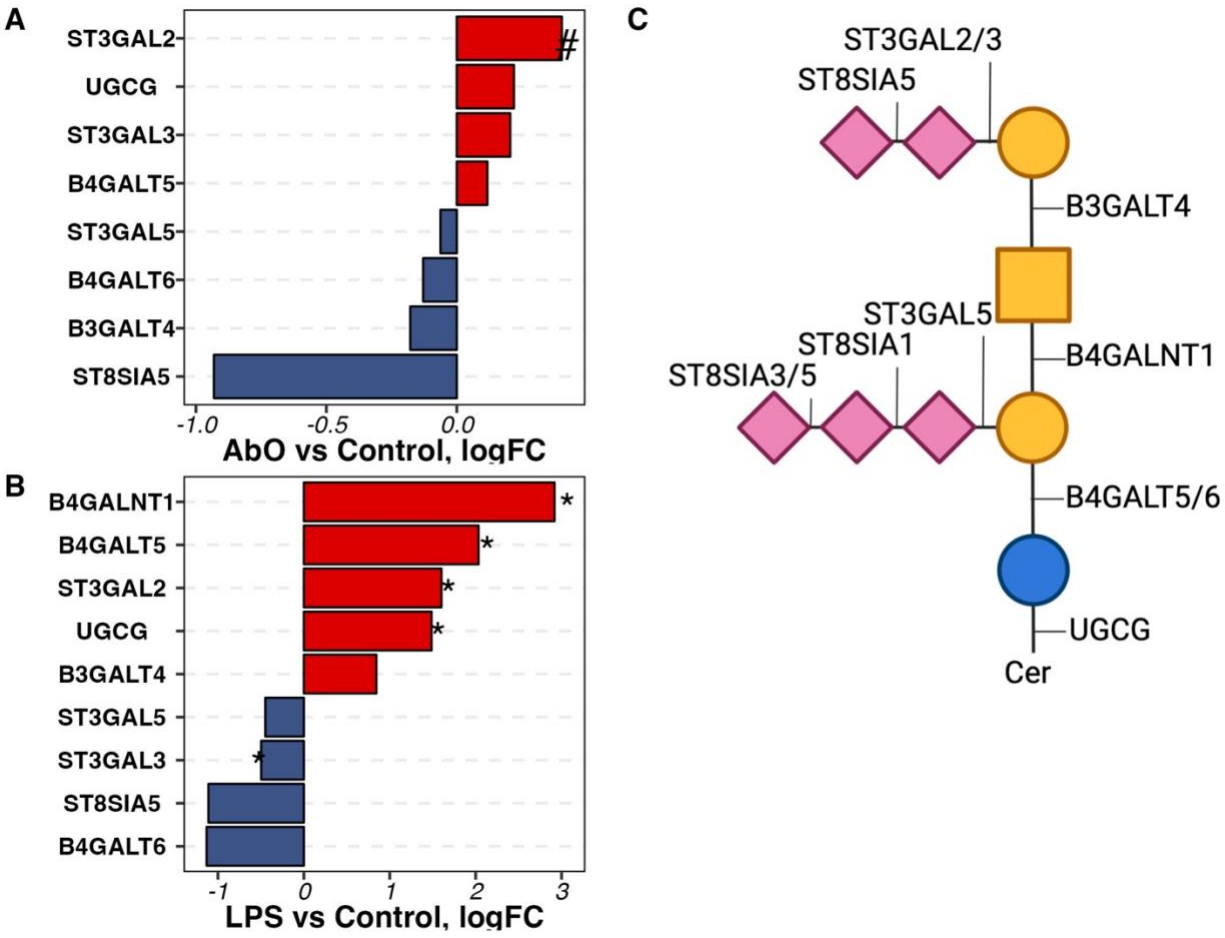
Figure 2-6 Summary of glycosylation and associated gene expression changes in hiMG stimulated by AβO and LPS. Representative examples of indicated glycan structures measured by mass spectrometry (MS) are listed in the middle column, and the genes linked to the indicated glycosylation changes identified by RNA-seq are listed in the right column. Note that current glycomic methods cannot precisely distinguish between core fucose and terminal fucose, and that the fucose residues in the structures such as C-FS_6513 can be either at the core or the terminal position. *The increased core fucosylation following AβO treatment is based on Jin et al., 2023.



Supplementary figure 2-S1 The missingness of N-glycans and GSL as the percentage of the number of samples. **A**) The missingness in N-glycan. **B**) The missingness in GSL. Red dash lines denote missingness of 80% across all samples.



Supplementary figure 2-S2 The scatter plot of the post-hoc linear discriminant analysis (LDA) of MANOVA for the PCA plots. **A**) The LDA scatter plot for N-glycan. According to the LDA scatter plot, LPS had the most significant impact on the group mean differences. **B**) The LDA scatter plot for GSL.



Supplementary figure 2-S3 The differential expression of genes involved in GSL biosynthesis in in AβO- and LPS-activated hiMG. **A**) The differential expression of genes involved in GSL biosynthesis in AβO-activated hiMG. **B**) The differential expression of genes involved in GSL biosynthesis in LPS-activated hiMG. **C**) The schematics of GSL structure and the corresponding enzymes. The pound (#) denotes genes with unadjusted p-value < 0.05 but FDR >0.05. The asterisk (*) denotes genes with FDR < 0.05.

REFERENCES

[1] Nimmerjahn A, Kirchhoff F, Helmchen F. Resting microglial cells are highly dynamic surveillants of brain parenchyma in vivo. Science (80-) 2005;308:1314–8. https://doi.org/10.1126/science.1110647.

[2] Streit WJ, Kreutzberg GW. Lectin binding by resting and reactive microglia. J Neurocytol 1987;16:249–60. https://doi.org/10.1007/BF01795308.

[3] Rebelo AL, Chevalier MT, Russo L, Pandit A. Role and therapeutic implications of protein glycosylation in neuroinflammation. Trends Mol Med 2022;28:270–89. https://doi.org/10.1016/j.molmed.2022.01.004.

[4] Marschallinger J, Iram T, Zardeneta M, Lee SE, Lehallier B, Haney MS, et al. Lipid-droplet-accumulating microglia represent a dysfunctional and proinflammatory state in the aging brain. Nat Neurosci 2020;23:194–208. https://doi.org/10.1038/s41593-019-0566-1.

[5] Bellenguez C, Küçükali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet 2022;54:412–36. https://doi.org/10.1038/s41588-022-01024-z.

[6] Johnson ECB, Dammer EB, Duong DM, Ping L, Zhou M, Yin L, et al. Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. Nat Med 2020;26:769–80. https://doi.org/10.1038/s41591-020-0815-6.

[7] Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 2019;51:404–13. https://doi.org/10.1038/s41588-018-0311-9.

[8] Heneka MT, Carson MJ, Khoury J El, Landreth GE, Brosseron F, Feinstein DL, et al. Neuroinflammation in Alzheimer's disease. Lancet Neurol 2015;14:388–405. https://doi.org/10.1016/S1474-4422(15)70016-5.

[9] Maezawa I, Nguyen HM, Di Lucente J, Jenkins DP, Singh V, Hilt S, et al. K v 1.3 inhibition as a potential microglia-Targeted therapy for Alzheimer's disease: Preclinical proof of concept. Brain 2018;141:596–612. https://doi.org/10.1093/brain/awx346.

[10] Maezawa I, Zimin PI, Wulff H, Jin LW. Amyloid-β protein oligomer at low nanomolar concentrations activates microglia and induces microglial neurotoxicity. J Biol Chem 2011;286:3693–706. https://doi.org/10.1074/jbc.M110.135244.

[11] Leng F, Edison P. Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here? Nat Rev Neurol 2021;17:157–72. https://doi.org/10.1038/s41582-020-00435-y.

[12] Penney J, Ralvenius WT, Tsai LH. Modeling Alzheimer's disease with iPSC-derived brain cells. Mol Psychiatry 2020;25:148–67. https://doi.org/10.1038/s41380-019-0468-3.

[13] Smith AM, Dragunow M. The human side of microglia. Trends Neurosci 2014;37:125–35. https://doi.org/10.1016/j.tins.2013.12.001.

[14] Saito H, Kashida S, Inoue T, Shiba K. The role of peptide motifs in the evolution of a protein network. Nucleic Acids Res 2007;35:6357–66. https://doi.org/10.1093/nar/gkm692.

[15] Gagneux P, Panin V, Hennet T, Aebi M, Varki A. Evolution of Glycan Diversity. 2022.

[16] Jin L-W, Di Lucente J, Ruiz Mendiola U, Tang X, Zivkovic AM, Lebrilla CB, et al. The

role of FUT8-catalyzed core fucosylation in Alzheimer's amyloid-β oligomer-induced activation of human microglia. Glia 2023;71:1346–59.

[17] Haenseler W, Sansom SN, Buchrieser J, Newey SE, Moore CS, Nicholls FJ, et al. A Highly Efficient Human Pluripotent Stem Cell Microglia Model Displays a Neuronal-Co-culture-Specific Expression Profile and Inflammatory Response. Stem Cell Reports 2017;8:1727–42. https://doi.org/10.1016/j.stemcr.2017.05.017.

[18] Lambert MP, Barlow AK, Chromy BA, Edwards C, Freed R, Liosatos M, et al. Diffusible, nonfibrillar ligands derived from Aβ1-42 are potent central nervous system neurotoxins. Proc Natl Acad Sci U S A 1998;95:6448–53. https://doi.org/10.1073/pnas.95.11.6448.

[19] Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 2019;37:907–15. https://doi.org/10.1038/s41587-019-0201-4.

[20] Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923–30. https://doi.org/10.1093/bioinformatics/btt656.

[21] McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 2012;40:4288–97. https://doi.org/10.1093/nar/gks042.

[22] Li Q, Xie Y, Wong M, Barboza M, Lebrilla CB. Comprehensive structural glycomic characterization of the glycocalyxes of cells and tissues. Nat Protoc 2020;15:2668–704. https://doi.org/10.1038/s41596-020-0350-4.

[23] Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Softw 2017;82:1–26. https://doi.org/10.18637/JSS.V082.I13.

[24] Gaunitz S, Tjernberg LO, Schedin-Weiss S. The N-glycan profile in cortex and hippocampus is altered in Alzheimer disease. J Neurochem 2021;159:292–304. https://doi.org/10.1111/jnc.15202.

[25] Li J, Hsu HC, Mountz JD, Allen JG. Unmasking Fucosylation: from Cell Adhesion to Immune System Regulation and Diseases. Cell Chem Biol 2018;25:499–512. https://doi.org/10.1016/j.chembiol.2018.02.005.

[26] Lünemann JD, von Gunten S, Neumann H. Targeting sialylation to treat central nervous system diseases. Trends Pharmacol Sci 2021;42:998–1008. https://doi.org/10.1016/j.tips.2021.09.002.

[27] Miwa HE, Song Y, Alvarez R, Cummings RD, Stanley P. The bisecting GlcNAc in cell growth control and tumor progression. Glycoconj J 2012;29:609–18. https://doi.org/10.1007/s10719-012-9373-6.

[28] Sandhoff R, Sandhoff K. Emerging concepts of ganglioside metabolism. FEBS Lett 2018;592:3835–64.

[29] Belarbi K, Cuvelier E, Bonte M-A, Desplanque M, Gressier B, Devos D, et al. Glycosphingolipids and neuroinflammation in Parkinson's disease. Mol Neurodegener 2020;15:59.

[30] Chen Q, Tan Z, Guan F, Ren Y. The Essential Functions and Detection of Bisecting GlcNAc in Cell Biology. Front Chem 2020;8. https://doi.org/10.3389/fchem.2020.00511.

[31] Park D, Brune KA, Mitra A, Marusina AI, Maverakis E, Lebrilla CB. Characteristic changes in cell surface glycosylation accompany intestinal epithelial cell (IEC) differentiation: High mannose structures dominate the cell surface glycome of undifferentiated enterocytes. Mol Cell Proteomics 2015;14:2910–21. https://doi.org/10.1074/mcp.M115.053983.

[32] Hinneburg H, Pedersen JL, Bokil NJ, Pralow A, Schirmeister F, Kawahara R, et al. High-resolution longitudinal N- And O-glycoprofiling of human monocyte-to-macrophage transition. Glycobiology 2020;30:679–94. https://doi.org/10.1093/glycob/cwaa020.

[33] Delannoy CP, Rombouts Y, Groux-Degroote S, Holst S, Coddeville B, Harduin-Lepers A, et al. Glycosylation Changes Triggered by the Differentiation of Monocytic THP-1 Cell Line into Macrophages. J Proteome Res 2017;16:156–69. https://doi.org/10.1021/acs.jproteome.6b00161.

[34] Fujitani N, Furukawa J ichi, Araki K, Fujioka T, Takegawa Y, Piao J, et al. Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. Proc Natl Acad Sci U S A 2013;110:2105–10. https://doi.org/10.1073/pnas.1214233110.

[35] Pace A, Scirocchi F, Napoletano C, Zizzari IG, D'angelo L, Santoro A, et al. Glycan-Lectin Interactions as Novel Immunosuppression Drivers in Glioblastoma. Int J Mol Sci 2022;23. https://doi.org/10.3390/ijms23116312.

[36] Haukedal H, Freude KK. Implications of Glycosylation in Alzheimer's Disease. Front Neurosci 2021;14. https://doi.org/10.3389/fnins.2020.625348.

[37] Akasaka-Manya K, Manya H, Sakurai Y, Wojczyk BS, Kozutsumi Y, Saito Y, et al.

Protective effect of N-glycan bisecting GlcNAc residues on β-amyloid production in Alzheimer's disease. Glycobiology 2010;20:99–106. https://doi.org/10.1093/glycob/cwp152.

[38] Gizaw ST, Ohashi T, Tanaka M, Hinou H, Nishimura SI. Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. Biochim Biophys Acta - Gen Subj 2016;1860:1716–27. https://doi.org/10.1016/j.bbagen.2016.03.009.

[39] Palmigiano A, Barone R, Sturiale L, Sanfilippo C, Bua RO, Romeo DA, et al. CSF N-glycoproteomics for early diagnosis in Alzheimer's disease. J Proteomics 2016;131:29–37. https://doi.org/10.1016/j.jprot.2015.10.006.

[40] Akasaka-Manya K, Manya H, Sakurai Y, Wojczyk BS, Spitalnik SL, Endo T. Increased bisecting and core-fucosylated N-glycans on mutant human amyloid precursor proteins. Glycoconj J 2008;25:775–86. https://doi.org/10.1007/s10719-008-9140-x.

[41] Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. Nat Rev Mol Cell Biol 2020;21:729–49. https://doi.org/10.1038/s41580-020-00294-x.

[42] Estus S, Shaw BC, Devanney N, Katsumata Y, Press EE, Fardo DW. Evaluation of CD33 as a genetic risk factor for Alzheimer's disease. Acta Neuropathol 2019;138:187–99. https://doi.org/10.1007/s00401-019-02000-4.

[43] Klaus C, Hansen JN, Ginolhac A, Gérard D, Gnanapragassam VS, Horstkorte R, et al. Reduced sialylation triggers homeostatic synapse and neuronal loss in middle-aged mice. Neurobiol Aging 2020;88:91–107. https://doi.org/10.1016/j.neurobiolaging.2020.01.008.

[44] Allendorf DH, Puigdellívol M, Brown GC. Activated microglia desialylate their surface, stimulating complement receptor 3-mediated phagocytosis of neurons. Glia 2020;68:989–98.

[45] Bhattacherjee A, Rodrigues E, Jung J, Luzentales-Simpson M, Enterina JR, Galleguillos D, et al. Repression of phagocytosis by human CD33 is not conserved with mouse CD33. Commun Biol 2019;2. https://doi.org/10.1038/s42003-019-0698-6.

[46] Pluvinage J V., Haney MS, Smith BAH, Sun J, Iram T, Bonanno L, et al. CD22 blockade restores homeostatic microglial phagocytosis in ageing brains. Nature 2019;568:187–92. https://doi.org/10.1038/s41586-019-1088-4.

[47] Allendorf DH, Franssen EH, Brown GC. Lipopolysaccharide activates microglia via neuraminidase 1 desialylation of Toll-like Receptor 4. J Neurochem 2020;155:403–16. https://doi.org/10.1111/jnc.15024.

[48] Gonzalez-Gil A, Schnaar RL. Siglec ligands. Cells 2021;10. https://doi.org/10.3390/cells10051260.

[49] Büll C, Nason R, Sun L, Van Coillie J, Sørensen DM, Moons SJ, et al. Probing the binding specificities of human Siglecs by cell-based glycan arrays. Proc Natl Acad Sci U S A 2021;118:e2026102118. https://doi.org/10.1073/pnas.2026102118.

[50] Gonzalez-Gil A, Porell RN, Fernandes SM, Maenpaa E, Li TA, Li T, et al. Human brain sialoglycan ligand for CD33, a microglial inhibitory Siglec implicated in Alzheimer's disease. J Biol Chem 2022;298. https://doi.org/10.1016/j.jbc.2022.101960.

[51] Brinkman-Van Der Linden ECM, Varki A. New aspects of siglec binding specificities,

including the significance of fucosylation and of the sialyl-Tn epitope. J Biol Chem 2000;275:8625–32. https://doi.org/10.1074/jbc.275.12.8625.

[52] Stansley B, Post J, Hensley K. A comparative review of cell culture systems for the study of microglial biology in Alzheimer's disease. J Neuroinflammation 2012;9. https://doi.org/10.1186/1742-2094-9-115.

[53] Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. EMBO Mol Med 2016;8:595–608. https://doi.org/10.15252/emmm.201606210.

[54] Suttapitugsakul S, Stavenhagen K, Donskaya S, Bennett DA, Mealer RG, Seyfried NT, et al. Glycoproteomics Landscape of Asymptomatic and Symptomatic Human Alzheimer's Disease Brain. Mol Cell Proteomics 2022;21:100433. https://doi.org/10.1016/j.mcpro.2022.100433.

[55] Yoshimura M, Ihara Y, Ohnishi A, Ijuhin N, Nishiura T, Kanakura Y, et al. Bisecting N-acetylglucosamine on K562 cells suppresses natural killer cytotoxicity and promotes spleen colonization. Cancer Res 1996;56:412–8.

[56] Tan Z, Wang C, Li X, Guan F. Bisecting N-acetylglucosamine structures inhibit hypoxia-induced epithelial-mesenchymal transition in breast cancer cells. Front Physiol 2018;9:210.

[57] de-Freitas-Junior JCM, Carvalho S, Dias AM, Oliveira P, Cabral J, Seruca R, et al. Insulin/IGF-I signaling pathways enhances tumor cell invasion through bisecting GlcNAc N-glycans modulation. an interplay with E-cadherin. PLoS One 2013;8:e81579.

[58] Kizuka Y, Kitazume S, Fujinawa R, Saito T, Iwata N, Saido TC, et al. An aberrant sugar

modification of BACE1 blocks its lysosomal targeting in Alzheimer's disease. EMBO Mol Med 2015;7:175–89. https://doi.org/10.15252/emmm.201404438.

[59] Sipione S, Monyror J, Galleguillos D, Steinberg N, Kadam V. Gangliosides in the Brain: Physiology, Pathophysiology and Therapeutic Applications. Front Neurosci 2020;14. https://doi.org/10.3389/fnins.2020.572965.

[60] Galleguillos D, Wang Q, Steinberg N, Zaidi A, Shrivastava G, Dhami K, et al. Anti-inflammatory role of GM1 and other gangliosides on microglia. J Neuroinflammation 2022;19. https://doi.org/10.1186/s12974-021-02374-x.

[61] Simon BM, Malisan F, Testi R, Nicotera P, Leist M. Disialoganglioside GD3 is released by microglia and induces oligodendrocyte apoptosis. Cell Death Differ 2002;9:758–67. https://doi.org/10.1038/sj.cdd.4401027.

[62] Abud EM, Ramirez RN, Martinez ES, Healy LM, Nguyen CHH, Newman SA, et al. iPSC-Derived Human Microglia-like Cells to Study Neurological Diseases. Neuron 2017;94:278-293.e9. https://doi.org/10.1016/j.neuron.2017.03.042.

[63] Guneykaya D, Ivanov A, Hernandez DP, Haage V, Wojtas B, Meyer N, et al. Transcriptional and Translational Differences of Microglia from Male and Female Brains. Cell Rep 2018;24:2773-2783.e6. https://doi.org/10.1016/j.celrep.2018.08.001.

[64] Grabert K, Michoel T, Karavolos MH, Clohisey S, Kenneth Baillie J, Stevens MP, et al. Microglial brain region-dependent diversity and selective regional sensitivities to aging. Nat Neurosci 2016;19:504–16. https://doi.org/10.1038/nn.4222.

Chapter 3: Glycosylation alterations in serum of Alzheimer's disease patients show widespread changes in N- glycosylation of proteins related to immune function, inflammation, and lipoprotein metabolism

Jennyfer Tena[1*], Xinyu Tang[2*], Qingwen Zhou[1], Danielle Harvey[3], Maria Barajas-Mendoza[1], Lee-Way Jin[3,4], Izumi Maezawa[3,4], Angela M. Zivkovic[2], Carlito B. Lebrilla[1**]

[1]Department of Chemistry, University of California, Davis, Davis, CA, USA

[2]Department of Nutrition, University of California, Davis, Davis, CA, USA

[3]Department of Pathology and Laboratory Medicine, School of Medicine, University of California, Davis, Sacramento, CA, USA

[4]UC Davis MIND Institute, Sacramento, CA, USA

*These two authors have equal contribution

**Corresponding author: Carlito B. Lebrilla, e-mail: cblebrilla@ucdavis.edu , One Shields Ave, Department of Chemistry, University of California, Davis, Davis, CA, USA

Published manuscript with links to all supplemental files available at: https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/dad2.12309

ABSTRACT

INTRODUCTION: There is an increased need for the development of novel blood-based biomarkers for early detection, prevention, or intervention in Alzheimer's Disease (AD). This study sought to determine whether serum glycopeptide analysis holds potential for identifying novel diagnostics and prognostics of AD.

METHODS: The study involved 195 participants including 96 patients with an AD diagnosis and 99 controls with no cognitive deficit. Utilizing a validated analytical mass spectrometry method, we monitored the site-specific glycosylation of 52 serum glycoproteins.

RESULTS: Partial least squares discriminant analysis revealed changes in overall sialylation and fucosylation of serum glycoproteins may be indicators of AD disease state. Loss of fucosylation of immunoglobulin G1 (IgG1) and IgG2 were indicative of AD diagnosis. Individual glycopeptide analysis found separation between the AD patients and controls on complement proteins and apolipoprotein B.

DISCUSSION: The results of this study suggest that serum glycoprofiling may be a promising approach for biomarker discovery.

Key words: Site-specific Glycosylation, Blood-based biomarker, Mass spectrometry, Alzheimer's Disease

INTRODUCTION

Alzheimer's disease (AD) is a chronic neurodegenerative disease that begins decades before the onset of symptoms and discernible dementia [1]. Medications for the treatment for AD in the later stages of disease have been elusive, with even the newest and most promising approaches only being able to slow the progression of disease but not prevent or reverse it [1]. Given the hallmark characteristics of AD, namely amyloid-beta (Aβ) deposition, neurofibrillary tangles involving hyperphosphorylated tau, and neurodegeneration, the current techniques for AD detection rely on cerebrospinal fluid (CSF) and plasma biomarkers (e.g. Aβ42/Aβ40 ratio) and images from positron emission tomography (PET) to evaluate the extent of neurodegeneration in the brain [2–4]. However, CSF-based biomarkers require invasive sample collection, the PET imaging-based detection is costly for screening purposes, and these detection approaches are generally uninformative about how to intervene to prevent or slow the progression of the disease. Thus, early, actionable detection approaches that could lead to the development of new disease prevention strategies are urgently needed.

Glycosylation is a common post-translational modification that affects protein structure and function with diverse glycans. Multiple studies have revealed aberrant glycosylation in AD-related proteins, including amyloid precursor protein (APP), tau, Beta-secretase 1 (BACE1), and Nicastrin (NCSTN)[5–8]. Furthermore, dysregulated glycosylation in AD brains affects multiple biological processes, such as neuroinflammation, cell adhesion, and cell signaling [8]. Plasma glycan-based measurements have proven useful for the detection of other diseases, including ovarian cancer and breast cancer [9,10].

Multiple reaction monitoring (MRM) technology has enabled large-scale glycoproteomic profiling of blood. Here we used an established MRM method [11,12] to evaluate the potential

for plasma or serum based glycoproteomic profiling as a viable biomarker approach for AD. We used serum samples obtained from 100 AD patients and 100 age- and gender matched cognitively normal controls whose samples are part of the UC Davis Alzheimer's Disease Research Center biorepository, to quantify serum glycopeptides of the most abundant serum proteins, and determine the potential of serum glycoprofiling as a novel biomarker approach for AD.

MATERIALS AND METHODS

Study Design

A power calculation was conducted on an exploratory set of samples involving 48 serum samples. Sample size calculation was performed prior to the study design as multiple two-sample T-tests with an FDR-adjusted rate of 0.05, and a K factor of 22. A sample size of 100 per group was determined to be an adequate number of participants for a power of 80%. In this study, samples from 195 total participants were analyzed, which included 99 participants with no cognitive deficit as per clinical diagnosis, and 96 participants with an AD diagnosis. All AD patients had a clinical AD diagnosis made within one year of the blood draw based on the UC Davis Alzheimer's Disease Research Center (ADRC) clinical diagnosis criteria. The diagnosis criteria for AD involved autopsy confirmation with BRAAK state four or higher with likelihood of AD moderate to high, with those patients for whom autopsy confirmation was not available being diagnosed with 'probable AD' based on the ADRC clinical diagnosis criteria. All human subjects provided informed consent as per the ADRC guidelines. The cohorts were selected to have equal percentages of males and females in both groups. In the control group 50% of participants were male while in the AD group 51% were male. **Table 3-1** summarizes the demographic data of all participants including age, sex, BMI, APOE genotype, and ethnicity, among other parameters. Among AD patients, 13.3% were African American, 10% were Hispanic, and 77% were White. Among controls, 20% were African American, 27% were Hispanic, and 53% were White. The median age was 79 [IQR: 73-83] for AD patients and 78 [IQR:73-82] for the normal controls. The median body mass index (BMI) was 26.4 [IQR: 23.83-29.64] for AD patients and 27.70 [IQR: 25.34-31.01] for the normal controls.

Sample Preparation

Serum samples were processed using a well-established protocol coupled to the dynamic multiple-reaction monitoring (dMRM) analytical method developed in our laboratory [13]. Briefly, serum samples were reduced with dithiothreitol, alkylated with iodoacetamide and digested with trypsin in a water bath at 37˚C for 18 h. For glycopeptide quantitative analysis, tryptic digested samples were analyzed directly with no further enrichment, as shown in **Figure 3-S1**. Serum samples were then spiked with a synthetic peptide standard for inter-batch reproducibility monitoring.

UPLC−ESI-QqQ−MS Analysis of Serum Glycoproteomic

The mixed samples were analyzed using ultrahigh-performance liquid chromatography (UPLC) coupled with triple quadrupole mass spectrometry as previously described [12]. Briefly, 2 μL samples were injected and separated by an Agilent Eclipse plus C18 column (rapid resolution high definition (RRHD) 1.8μm, 2.1×100 mm) coupled with a C18 guard column (RRHD 1.8μm, 2.1×5 mm). An Agilent 1290 infinity LC system (Agilent Technologies, Santa Clara, CA) was used, and the separation was performed with a 70 min binary gradient consisting of solvent A of 3% acetonitrile, 0.1% formic acid, solvent B of 90% acetonitrile, and 0.1% formic acid in nano pure water (v/v) at a flow rate of 0.5mL/min. The UPLC system was coupled to an Agilent 6490 triple quadrupole (QQQ) mass spectrometer (Agilent Technologies, Santa Clara, CA) with the MS conditions previously described [11]. The dMRM method used predetermined collision induced dissociation from previously determined study however LC retention times were adjusted. The list of the serum glycoproteins monitored are shown in **Table 3-S1**, which involved over 400 transitions. A deep learning method PB-Net (Peak Boundary Neural Network) was used for fully automatic chromatographic peak integration [14]. Relative

glycopeptide abundances were calculated using the area under the curve of the glycopeptide and normalized to its reference non-glycosylated peptide as previously described [13].

Data Processing and Statistical Analysis

Instrument reproducibility was monitored with commercially available digested serum samples (Sigma-Aldrich) as quality controls. The digested serum pool was used as a quality control for monitoring MRM transitions every ten samples for a total of 20 technical replicates. A hierarchical clustering was then performed using the relative abundance of all glycopeptides with average linkage to detect any outlier samples. Samples were determined to be outliers at a branch cutoff at height 15.5 using hierarchical clustering and were removed from analysis. The coefficient of variation (CV) was calculated for all glycopeptides based on these pooled serum samples. Glycopeptides that were below the limit of quantitation for the instrument or with CV>30% were excluded from analysis. The naming convention of the glycopeptides used throughout the text follows the convention, Protein_Glycosite_Glycan. For example, IgG1_297_5510 reads Immunoglobulin G1 at glycosite N297 with N-glycan composition of $Hex_{(5)}HexNAc_{(5)}Fuc_{(1)}Neu5Ac_{(0)}$.

The fractions of sialylated and fucosylated peptides were calculated as the sum of relative glycopeptide abundances of non-, mono-, di-, or poly-glycosylated peptides relative to that of the total peptides. Relative glycopeptide abundances of individual glycopeptides and the fractions of sialylated and fucosylated peptides were used in subsequent analyses. All statistical analyses were performed using the R statistical package (R version 4.1.0). Linear models were constructed for univariate and multivariate analysis that included each potential covariate separately. Each simple multivariate model had one covariate at a time: glycopeptide abundance ~ Diagnosis + covariate. These simple multivariate models were compared to the univariate

model which included diagnosis group as the only variable to evaluate the impact of these covariates on glycopeptide concentrations. Potential covariates included age, sex, ethnicity, presence of the ApoE4 allele, and BMI. Linear models with interaction terms were used to analyze the sex-by-AD diagnosis interaction and ethnicity-by-AD diagnosis interaction: glycopeptide abundance ~ diagnosis + sex + diagnosis * sex or glycopeptide abundance ~ diagnosis + ethnicity + Diagnosis * ethnicity. For sex-by-AD diagnosis interaction, a post hoc analysis was done to estimate the sex specific diagnosis effects on serum glycopeptide abundance.

Differential expression analysis was performed as linear models using the limma package in R. Coefficients related to diagnosis were tested according to the null hypothesis (being zero) using t-tests moderated in a Bayesian fashion. Raw p-values were then adjusted for multiple hypothesis testing using the Benjamini-Hochberg (BH) correction. The differences in glycopeptide abundance between normal control and AD are presented as natural log fold-change.

A partial least squares discriminant analysis (PLS-DA) was performed to assess the discriminatory potential of serum glycopeptides. We scaled glycopeptide abundances to the variance of 1 and conducted leave-one-out cross-validation (LOOCV) to identify the best number of latent components using 1 through 20 with the train() function from the caret package. The importance of independent variables in the PLS-DA model was measured using variable importance in projection (VIP) scores. Spearman's correlation analysis was used to evaluate the association between glycopeptides, cognitive scores (Spanish and English Neuropsychological Assessment Scales (SENAS) scores and clinical dementia rating (CDR).

RESULTS

A hierarchical clustering was first performed to target outlier samples that can result in false positives. The hierarchical clustering results show three outlier samples that were excluded from all further analysis (B10, C13, and A67) as shown in **Figure 3-S2**. CVs were calculated for all 372 glycopeptides monitored in this study and their distribution is listed in **Figure 3-S3**. Of the 372 glycopeptides monitored 14 glycopeptides had CV values greater than 30% as listed in **Table 3-S2**.

Fucosylated and Sialylated Glycopeptides in AD

We grouped glycopeptides by their fucosylation and sialylation status separately and assessed whether the fucosylation and sialylation status of glycopeptides discriminates serum samples from AD patients and normal controls by looking at patterns of fucosylated and sialylated glycopeptides generated from PLS-DA analysis. The fucosylation and sialylation status of glycopeptides did not explicitly separate AD samples from normal controls with PLS-DA models (**Figure 3-1A and 3-1C**). However, the models show that several changes in overall sialylation and fucosylation of serum glycoproteins may be indicators of AD disease state. The PLS-DA model for fucosylation status had 12 glycopeptides with VIP scores > 1.5. A VIP score > 1.5 is considered to enable discrimination between groups. Notably, non-fucosylated and mono-fucosylated immunoglobulin G1 (IgG1) and IgG2 had high contributions to the model with VIP score > 2, indicating the fucosylation status of IgG1 and IgG2 may be potential biomarkers for AD (**Figure 3-1B**). The PLS-DA analysis for sialylation identified 19 glycopeptides with VIP scores > 1.5, which includes inflammation response glycoproteins such as Kininogen-1 (KNG1) and immune response glycoproteins such as Complement factor I

(CFAI) as well as lipid metabolism glycoproteins Apolipoprotein C3 (APOC3) and Clusterin (CLUS) (**Figure 3-1D**).

Identification of Aberrantly Fucosylated or Sialylated Glycopeptides in AD

The PLS-DA analysis for fucosylation and sialylation showed that AD and control samples were located in two clusters in the scores plot with some overlap. We then identified fucosylated and sialylated glycopeptides that drive the discrimination for AD versus normal controls using differential analysis. For fucosylation, six statistically significant glycoproteins which correspond to 10 glycopeptides differed (P-values < 0.05) between AD and normal controls in the univariate model, including glycopeptides with high contribution to the fucosylation PLS-DA model (non-fucosylated IgG1, monofucosylated IgG1, non-fucosylated IgG2, monofucosylated IgG2, etc.) (**Figure 3-1E-H and Figure 3-S4A**). For sialylation, eight proteins with 10 glycopeptides differed between AD and normal controls (**Figure 3-S4B**).

Glycopeptide Signatures of AD

We then used PLS-DA analysis to model the individual glycopeptide data and found separation between the AD patients and normal controls, though as with the fucosylation and sialylation data there was overlap between the groups (**Figure 3-2A**). The model had 51 glycopeptides that had a VIP score >1.5, including proteins involved in lipid metabolism (APOB, APOC3, APOH), immunity (IgG1, IgG2, IgM), and inflammation response (CO3, CFAI, VTNC, ANT) (**Figure 3-2B**). We then determined whether there were significant differences between patients and controls in the abundances of the glycopeptides that had high contribution to the PLS-DA model using differential analysis. CFAI_494_5402 and CO3_85_5200 were significantly increased and APOB_983_5401 was significantly decreased in the AD patients

compared to controls (P-values < 0.05). The three glycopeptides are all non-fucosylated (**Figure 3-2C-E**).

Identification of Aberrant Glycopeptides in AD

We next performed differential analysis to identify which glycopeptide abundances differed in the serum of AD patients vs. normal controls. With the univariate model, among the 372 glycopeptide transitions monitored accounting for 53 out of the 57 glycoproteins, 35 glycopeptide abundances were altered in individuals with AD, including 19 upregulated glycopeptides and 16 downregulated glycopeptides (P-values < 0.05, **Figure 3-3A**). Among the 35 glycopeptides that were differentially expressed in AD patients, most belong to proteins involved in immune function, including immunoglobulins and acute phase proteins such as alpha-1 antitrypsin (A1AT), alpha-2- macroglobulin (A2MG), alpha-1-acid-glycoprotein (AGP1), alpha-2-HS-glycoprotein (FETUA), and complement C3 (CO3). After adjusting for sex and age separately, these differences remain statistically significant. With the models adjusting for ethnicity, APOE4, and BMI separately the majority of the differentially expressed glycopeptides remained statistically significant, however, new differences were also detected in these multivariate models as shown in the x-axis of **Figure 3-3A**. In the final multivariate model, all covariates including sex, age, ethnicity, APOE4 status, and BMI were included. This analysis yielded 24 significant AD-associated glycopeptides (**Figure 3-3A**). The glycopeptides with highest VIP scores in the PLS-DA model, such as CFAI_494_5402, APOB_983_5401, A1AT_271MC_5402, FETUA_346_1101, etc., were also significantly different by differential expression analysis (P-values < 0.05), indicating they are important in distinguishing AD patients from controls.

Potential Confounders

Factors including sex, ethnicity, age, BMI, and ApoE4 genotype had impacts on glycopeptide abundances in serum. Sex and ethnicity were the two major confounders that influenced glycopeptide abundance in controls and AD. We analyzed the contributions of sex and ethnicity on serum glycopeptide abundances as covariates in the linear regression models. Among 372 glycopeptides, 135 glycopeptide abundances differed by sex, including 50 that were significantly higher in females and 85 that were significantly higher in males (adjusted P-values < 0.05, **Figure 3-S5 A**). Thirty-eight glycopeptides significantly differed across ethnicity groups, where KLKB1_127_5410 had the lowest P-value (adjusted P-values<0.05, **Figure 3-S5 B-D**). The impacts of other confounders are shown in **Figure 3-S5 E-G**.

Our study also provided the evidence of sex-by-AD diagnosis and ethnicity-by-AD diagnosis interactions. For example, ITIH4_517_5420/5401 was decreased in male AD patients vs. controls (P-value<0.05) but was not different in female AD patients vs. controls (**Figure 3-3B**). In contrast, IgG1_297_5510 was increased in AD female patients vs. controls (P-value<0.05) but was not different in male patients vs. controls (**Figure 3-3C**). For the ethnicity specific effects, Diagnosis had a different effect in glycopeptides such as IgA2 _201_5511, IgM _209_5511 depending on the ethnicity of subjects (**Figure 3-3, D&E**).

Associations Between Glycopeptides and Clinical Cognitive Scores

Correlation analysis revealed associations between cognitive scores and glycopeptides that were important in discriminating AD patients and controls. For example, non-fucosylated and monofucosylated IgG2, the former of which was increased and the latter of which was decreased in AD, were positively and negatively correlated with the CDR (**Figure 3-S6, A&B**), respectively. CO3_85_5200, which highly contributed to the PLS-DA model and which was

increased in AD patients vs. controls, was negatively associated with semantic score, and positively associated with CDR (**Figure 3-S6, C&D**).

DISCUSSION

In this study we set out to determine whether serum-based glycopeptide analysis may be a useful approach for identifying novel actionable diagnostics for AD. Our results suggest that serum glycopeptide profiling is a promising approach for the development of new biomarkers. Even with this platform, which is not yet tailored for AD-specific glycan markers, and instead is an application of an existing platform developed originally for cancer diagnostics, our results indicate that glycopeptide profiling can uncover important underlying disease mechanisms and point to future biomarker development. The importance of the fucosylation status of IgG in discriminating between controls and AD patients stood out as an important finding (**Figure 3-1B**). As the most abundant antibody in human blood, the importance of IgG in inflammation, infections, metabolic health and autoimmunity is well established [15,16]. Specifically, an increased abundance of non-fucosylated IgG1 and IgG2 was observed in AD patients compared with controls (**Figure 3-1, E&G**). Previous investigations have associated dementia with IgG N-glycans, however this previous work lacked site-specific information and did not include a diverse cohort in the population selection [17]. The extent of IgG fucosylation and sialylation has been found previously to be associated with the pro- vs. anti-inflammatory signaling of IgG, with non-fucosylated and non-sialylated IgG preferentially binding to pro-inflammatory Fcγ receptors [18]. Additionally, PLS-DA analysis of glycopeptides was able to separate AD patients from controls (**Figure 3-2A**) and proteins with VIP score >1.5 also involved immunoglobulins IgG1, IgG2, IgA and IgM among others (**Figure 3-2B**). Together, our results highlight the dysregulation of glycosylation of immune response proteins, particularly immunoglobulins, in AD pathology. Other glycopeptides with VIP score >1.5 included proteins involved in lipid metabolism (APOB, APOC3, APOH) and inflammation response (CO3, CFAI, VTNC, ANT)

(**Figure 3-2B**). For example, the mono-sialylated glycopeptide APOB_N983_5401 was decreased in AD patients compared to controls (**Figure 3-2D**). Previous studies have shown a relationship between the glycosylation state of lipoprotein-associated proteins and the pro-inflammatory capacity and functionality of their lipoprotein carrier [19,20].

The univariate model showed 35 altered glycopeptide abundances in AD patients from the 372 glycopeptide transitions monitored (**Figure 3-3A**). All altered proteins are involved in immune function and immune response, including immunoglobulins and acute phase proteins. Interestingly, all immunoglobulin glycopeptides were observed to be increased in AD patients. The study cohort involved equal numbers of males and females, and equal average age in both groups, thus when adjusting for sex and age all glycopeptide remained statistically significant as expected. Age and sex are the covariates that are typically accounted for in biomarker discovery, however, our study suggests that ethnicity is also an important variable that can contribute to variability (**Figure 3-3A, D & E**). Research on racial disparities in biomarkers for AD has shown that analyses of molecular biomarkers of AD should adjust for race. Given the low sample size for participants of different ethnic groups in this preliminary study we are not able to fully explore the contribution of ethnicity to the variability in serum glycopeptide profiles. However, our results indicate that serum glycoprofiling may be an exceptionally useful tool for the development of biomarkers in a disease such as AD, in which ethnic-specific differences in disease pathophysiology and biomarker profiles are marked. For instance, a differential effect on total tau and phosphorylated tau181 has been shown to differ in African American individuals compared to white individuals [21]. Our data revealed that certain glycopeptides such as KLKB1_127_5410 strongly distinguished between AD patients and controls only in African American participants but not in White or Hispanic participants, highlighting the potential for the

development of glycan-based diagnostics that have the ability to identify individuals at risk of developing AD across a broad diversity of individuals (**Figure 3-S5, B-D**).

Of note, our data revealed that non-fucosylated IgG2 was positively correlated with CDR while mono-fucosylated IgG2 was negatively correlated with CDR (**Figure 3-S6 A and B**). This result suggests a direct relationship between immunoglobulin fucosylation and clinical measures of cognitive function loss. Likewise, a non-sialylated, non-fucosylated glycopeptide of CO3 was also positively correlated with CDR (**Figure 3-S6 C**), highlighting the loss of fucosylation and sialylation at specific sites of immune-related proteins as a potential mediator and/or biomarker of AD. Major challenges in the development of blood-based biomarkers for AD include a high degree of patient heterogeneity and the potential existence of multiple clinical phenotypes [22–24]. Importantly, new biomarkers are needed not just for diagnostic purposes but also for determining prognosis and monitoring therapeutic efficacy. In this pilot work, we sought to determine whether serum-based glycoprofiling may be a useful approach to address these critical questions in the field. Our work for the first time utilized a previously developed MRM analytical method to monitor glycan alterations in serum glycoproteins in a site-specific manner to determine whether serum glycoprofiling can discriminate between AD patients and controls, but also simultaneously point to potential therapeutic approaches to prevent, treat, or reverse AD. Our results suggest that serum glycoprofiling may indeed be a powerful biomarker discovery platform that can be harnessed to identify new markers in diverse cohorts and new therapeutic targets, particularly those focused on immune-related mechanisms. Future studies including a larger sample size, particularly in ethnically diverse cohorts powered to account for age, sex, and ethnicity, are needed to develop glycan-based biomarkers for AD. Further work to refine the

transitions monitored by the MRM method that are specific to AD would improve the sensitivity and specificity of the biomarkers.

DECLARATIONS

Author contributions

Designed the study: L.W.J, A.M.Z, C.B.L, J.T, D.H

Supervised the study: A.M.Z, C.B.L

Performed Experiments: J.T, Q.Z

Performed statistical analysis: X.T, J.T, D.H

Analyzed data: J.T, Q.Z, M.B, X.T

Interpret data: A.M.Z, J.T, X.T, C.B.L

Wrote the manuscript: A.M.Z, X.T, C.B.L, J.T

All authors agree with the submission and have no conflicts of interest to disclose.

ACKNOWLEDGEMENTS

Research was supported by the National Institute on Aging of the National Institutes of Health under Award Number R01AG062240. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

TABLES AND FIGURES

Table 3-1: Demographics data of participants with Alzheimer's Disease and controls.

	Control	AD
n	99	96
Age (years) (median [IQR])	78[73-82]	79[78-83]
Gender=Male (%)	49 (49.5)	49 (51)
BMI (median [IQR])	27.7[25.3-31]	26.4[27-29.5]
APOE genotype		
e22 (%)	1(1.08)	0(0)
e23 (%)	13(14)	1(1.06)
e24 (%)	1(1.08)	3(3.19)
e33 (%)	55(59.1)	35(37.2)
e34 (%)	23(24.7)	42(44.7)
e44 (%)	0(0)	13(13.8)
APOE4 Positivity (%)	30 (30.3)	60 (62.5)
Ethnicity		
African American (%)	20(20.4)	12(12.5)
Hispanic (%)	26(26.5)	10(10.4)
White (%)	52(53.1)	74(77.1)
Verbal Memory Score (median [IQR])	0.5[-0.13-0.94]	-1.4[-1.3–1]
Executive Function Score (median [IQR])	0.15[-0.11-0.48]	-0.72[-0.63–0.47]
Semantic Memory Score (median [IQR])	0.66[-0.013-1.2]	-0.33[-0.42-0.17]
Spatial Score (median [IQR])	0.45[-0.11-0.82]	-0.51[-0.47-0.14]
CDR (median [IQR])	0[0-0.38]	4.5[5.3-7]
Total White Matter Hypertension (median [IQR])	5.7[2.7-13]	8.4[15-18]
Intracranial Volume (median [IQR])	1300[1200-1400]	1300[1300-1400]

Note:

Cognitive function scores and brain measurements had missing values in both groups. The number of data points for each measurement was listed in **supplemental table 3-S3**.

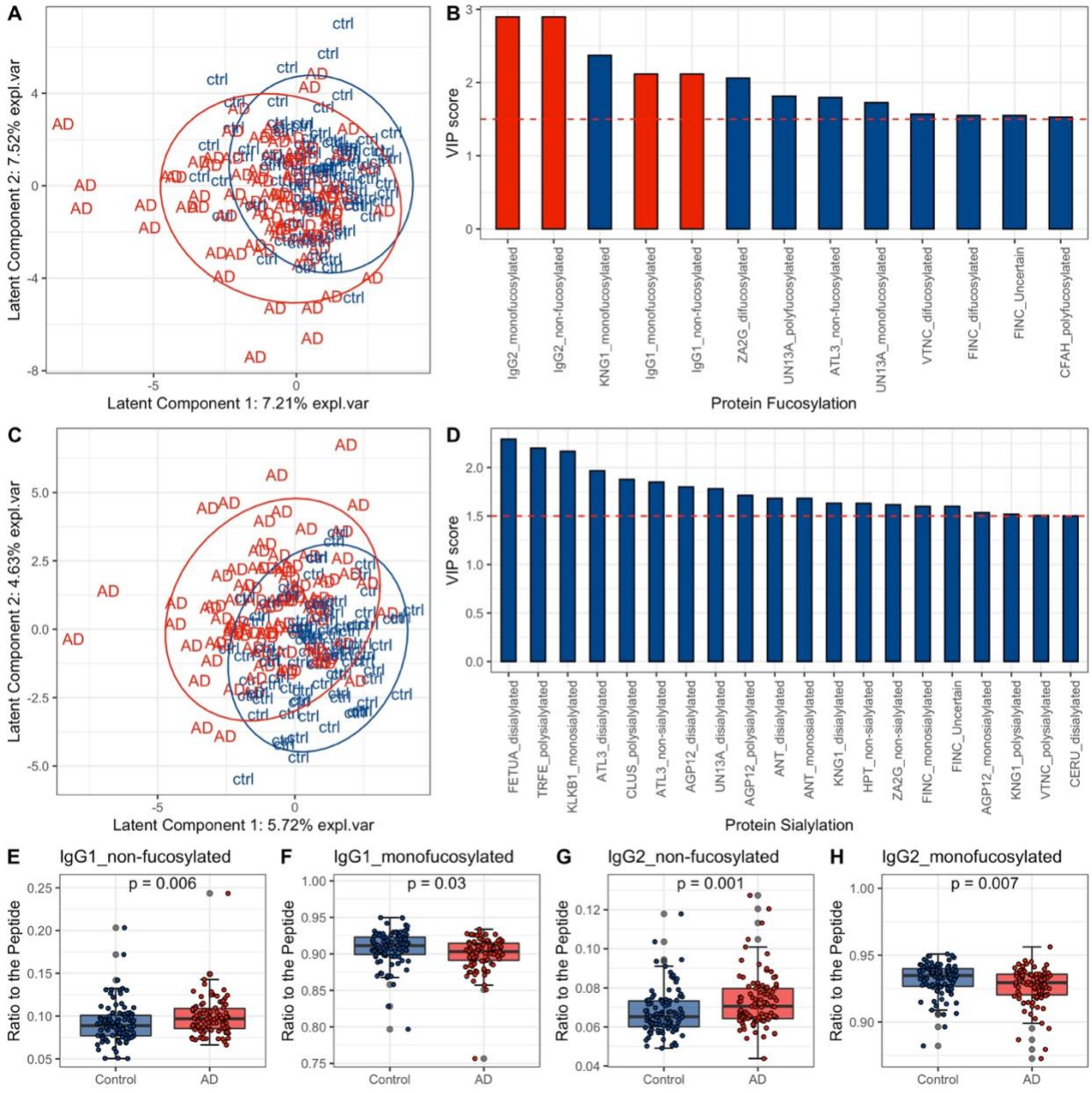


Figure 3-1: Changes in fucosylation and sialylation status in serum samples from patients with clinically diagnosed Alzheimer's Disease (AD) compared to controls (ctrl) using partial least squares-discriminant analysis (PLS-DA). **A)** The scores plot shows the distribution of subjects across latent components 1 and 2 given the PLS-DA model for fucosylation status. **B)** Variables with variable importance in projection (VIP) scores >= 1.5 and their VIP scores in the PLS-DA model for fucosylation status. Variables colored in red are those selected for differential analysis

and are visualized in E-H. **C)** The scores plot shows the distribution of participants across latent components 1 and 2 given the PLS-DA model for sialylation status. **D)** Variables with VIP scores >= 1.5 and their VIP scores in the PLS-DA model for sialylation status. **E-H)** Boxplots showing the differences in abundance of non-fucosylated and monofucosylated immunoglobulin G1 (IgG1) and IgG2 proteins from control and AD samples.

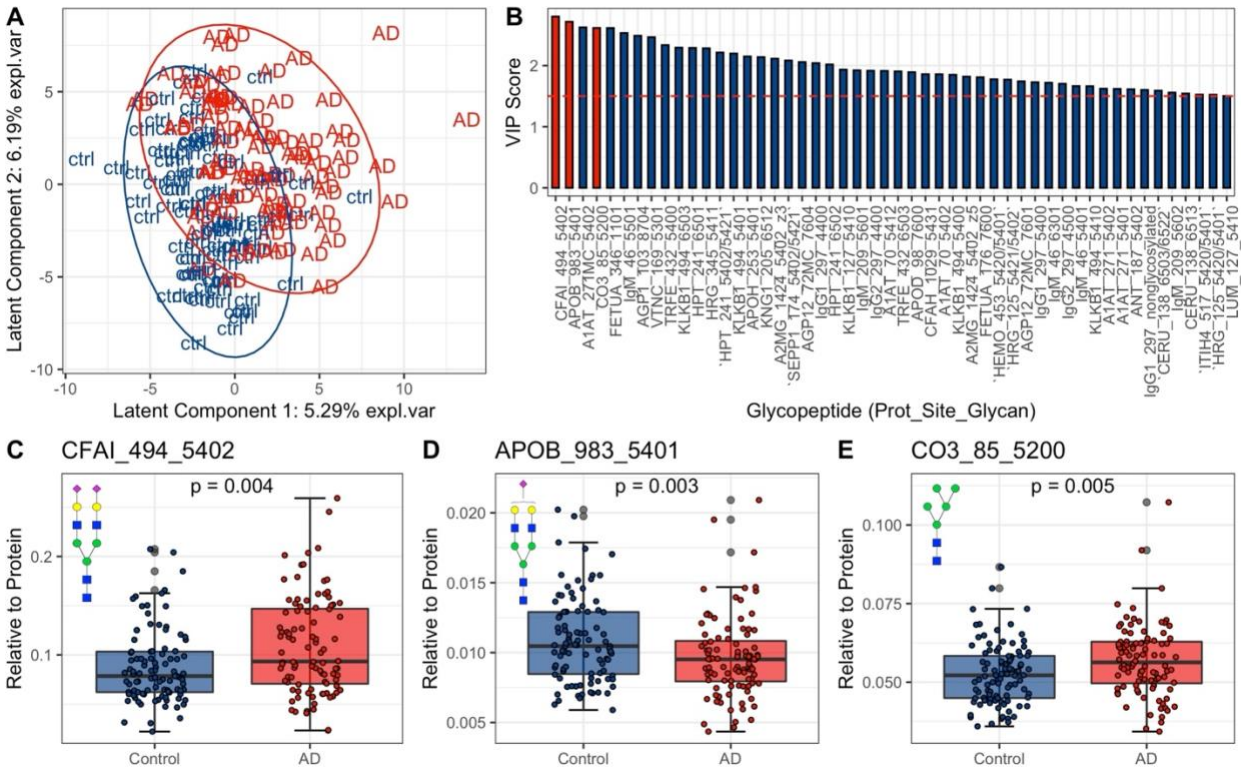


Figure 3-2: Changes in serum glycopeptide abundance in serum samples from Alzheimer's Disease patients compared to controls using PLS-DA. **A)** The scores plot shows the distribution of participants across latent components 1 and 2 given the PLS-DA model for glycopeptides. **B)** Variables with VIP scores >= 1.5 and their VIP scores. Variables colored in red are those selected for differential analysis and are visualized in C-E. **C-E)** Boxplots showing the differences in abundances of glycopeptides from control and Alzheimer's Disease (AD) samples. Protein abbreviations: complement factor I (CFAI), Apolipoprotein B (Apolipoprotein B) and Complement C3 (CO3). N-glycan symbol key: yellow circles, galactose (Gal); green circles, mannose (Man); blue squares, N-acetylglucosamine (GlcNAc); red triangles, fucose (Fuc); purple diamonds, N- acetylneuraminic acid (Neu5Ac).

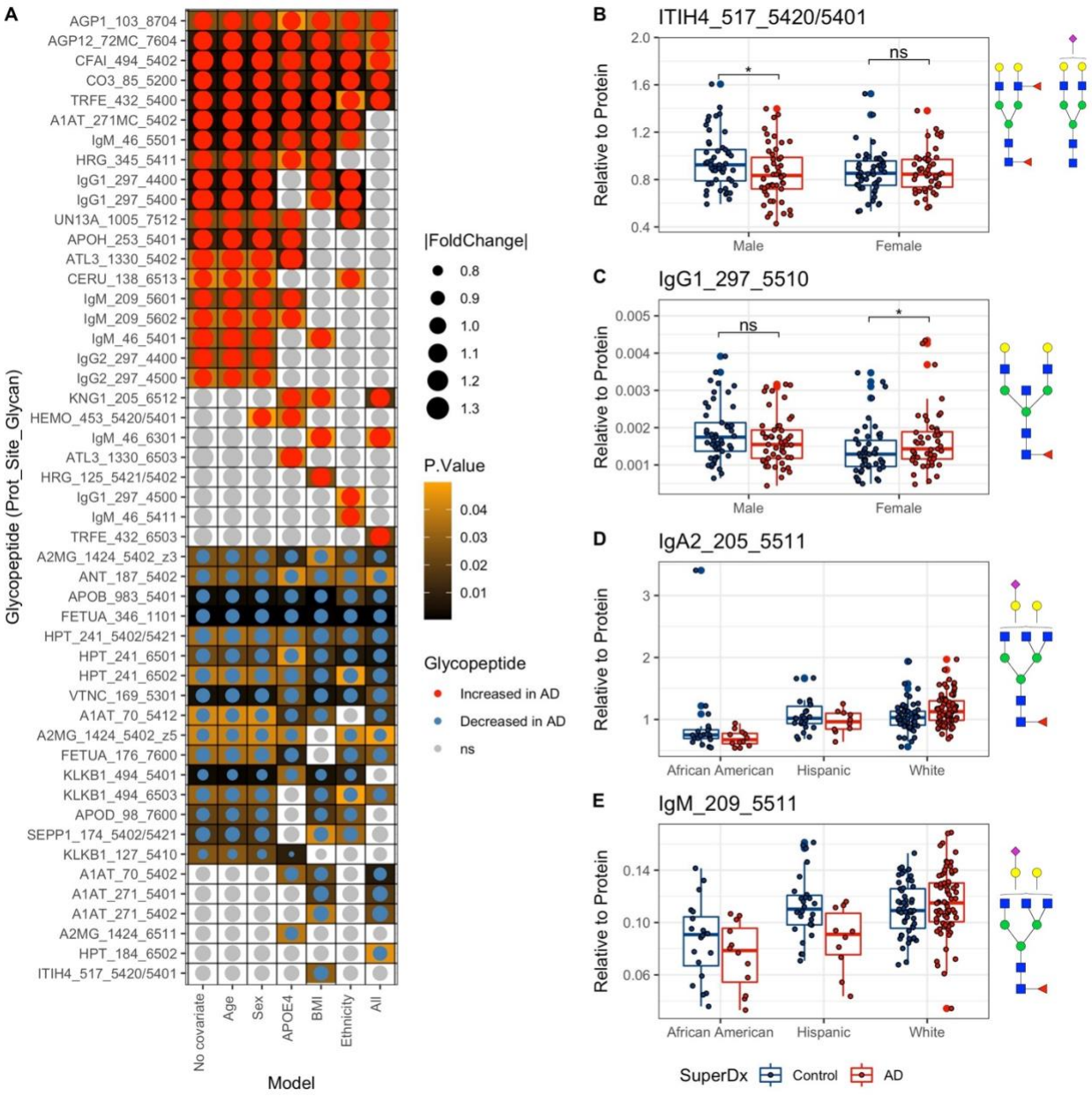


Figure 3-3: Differential analysis of glycopeptides and the impact of confounders on glycopeptide abundance. **A)** The heatmap shows glycopeptides significantly different in Alzheimer's Disease patients compared to controls using univariate and multivariate linear regressions. The dot size shows the effect size (absolute fold change). Darker background color refers to smaller p-value (before multiple testing correction). **B-C)** The boxplot shows the sex specific diagnosis effects on serum glycopeptide abundance. Differences with p-value < 0.05 are

denoted by asterisk, whereas differences with p>0.05 are denoted as “ns” (not significant). **D-E)** Boxplot showing the ethnicity specific diagnosis effects on serum glycopeptide abundance. Protein abbreviations: Inter-alpha-trypsin inhibitor heavy chain H1 (ITIH), Immunoglobulin G1 (IgG1), Immunoglobulin A2 (IgA2), and Immunoglobulin M (IgM). N-glycan symbol key: yellow circles, galactose (Gal); green circles, mannose (Man); blue squares, N-acetylglucosamine (GlcNAc); red triangles, fucose (Fuc); purple diamonds, N- acetylneuraminic acid (Neu5Ac).

Supplemental Table 3-S1: The MRM glycoproteins.

Glycoproteins
1_A1AT_Alpha-1-antrypsin
2_A1BG_Alpha-1B-glycoprotein
3_A2GL_Leucine-richAlpha-2-glycoprotein
4_A2MG_Alpha-2-macroglobulin
5_AACT_Alpha-1-antichymotrypsin
6_AFAM_Afamin
7_AGP1_Alpha-1-acid glycoprotein 1
8_AGP2_Alpha-1-acid glycoprotein 2
9_ANT_Antithromnin-III
10_APOB_ApolipopioteinB
11_APOC3_Apolipoprotein CIII
12_APOD_ApolipopioteinD
13_APOH_Beta-2-glycoprotein1
14_APOM_ApolipoproteinM
15_ATL3_3ADAMTS-likeProtein3
16_ATRN_Attractin
17_C1S_Complement C1s subcomponent
18_CAN3_Calpain-3
19_CERU_Ceruloplasmin
20_CFAH_ComplementFactorH
21_CFAI_ComplementFactorI
22_CLUS_Clusterin
23_CO2_ComplementC2
24_CO3_ComplementC3
25_CO4A_ComplementC4-A
26_CO4B_ComplementC4-B
27_CO6_ComplementComponetC6
28_CO8A_ComplementComponetC8AChain
29_CO8B_ComplementComponetC8BChain
30_FETUA_Alpha-2-HS-glycoprotein
31_FHR1_Complement factor H-related protein 1
32_FINC_Fibronectin
33_HEMO_Hemopexin
34_HEP2_HeparinCofactor2
35_HPT_Haptoglobin
36_HRG_Histidine-richGlycoprotein
37_IgA1_Immunoglobulin A1
38_IgA2_Immunoglobulin A2
39_IgG1_Immunoglobulin G1
Glycoproteins
40_IgG2_Immunoglobulin G2

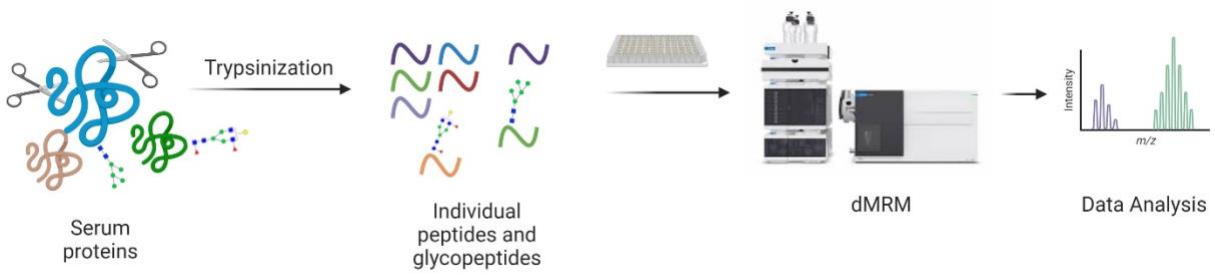
41_IgG3_Immunoglobulin G3
42_IgG4_Immunoglobulin G4
43_IgJ_Immunoglobulin J
44_IgM_Immunoglobulin M
45_ITIH1_Inter-alpha-trypsinInhibitorHeavyChainH1
46_ITIH4_Inter-alpha-trypsinInhibitorHeavyChainH4
47_KLKB1_PlasmaKallikrein
48_KNG1_Kininogen-1
49_KPCD3_Serine/threonine-proteinKinaseD3
50_LUM_Lumican
51_PON1_SerumParaoxonase/arylesterase1
52_SEPP1_SelenoproteinP
53_THBG_Thyroxine-bindingGlobulin
54_THRB_Prothrombin
55_TRFE_Transferrin
56_UN13A_Protein unc-13HomologA
57_VTNC_Vitronectin
58_ZA2G_Zinc-alpha-2-glycoprotein total

Supplemental Table 3-S2: The glycopeptides with CV over 30%.

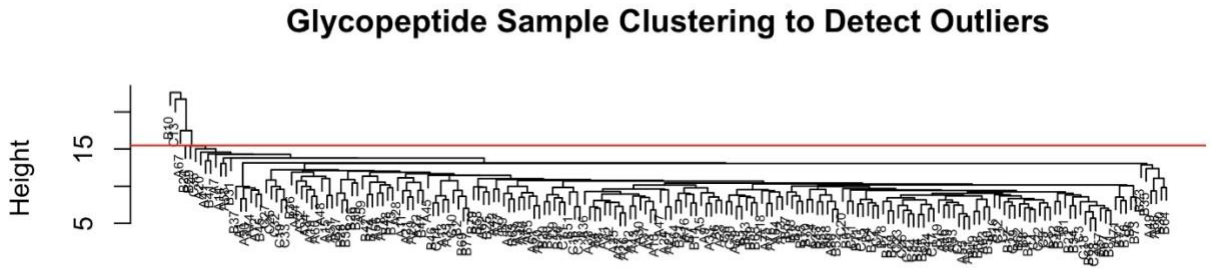
Glycopeptide	CV%
A1AT_70_5402	52
A1AT_70_5412	49
A2MG_247_5200	32
A2MG_247_5401_z4	35
A2MG_247_5402	37
AGP12_72MC_6503	34
AGP12_72MC_7604	35
APOC3_74_1210	57
APOC3_74_1300	47
CERU_397_6411	31
CFAI_494_5402	48
HPT_184_5400	31
HPT_241_5412	32
IgM_209_4601	38

Supplemental Table 3-S3: The number of data points for cognitive function measurements.

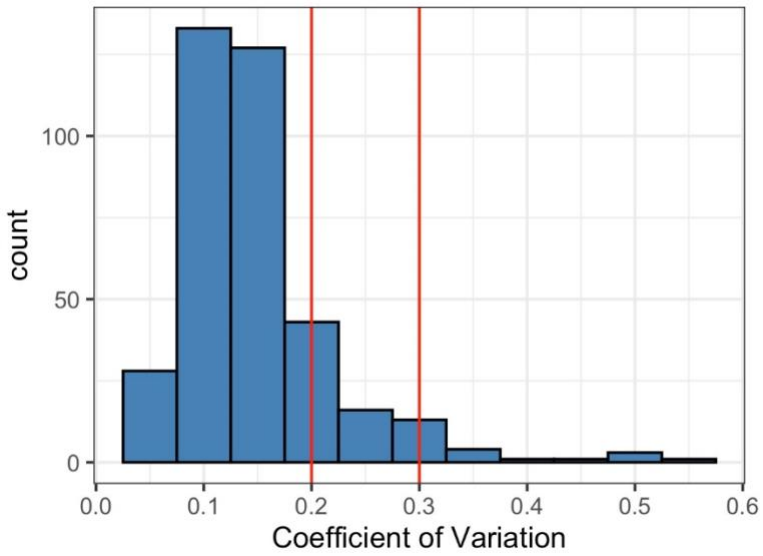
	Control	AD
Verbal Memory Score	84	50
Executive Function Score	84	54
Semantic Memory Score	70	54
Spatial Score	69	46
CDR	82	52
Total White Matter Hypertension	61	46
Intracranial Volume	61	46



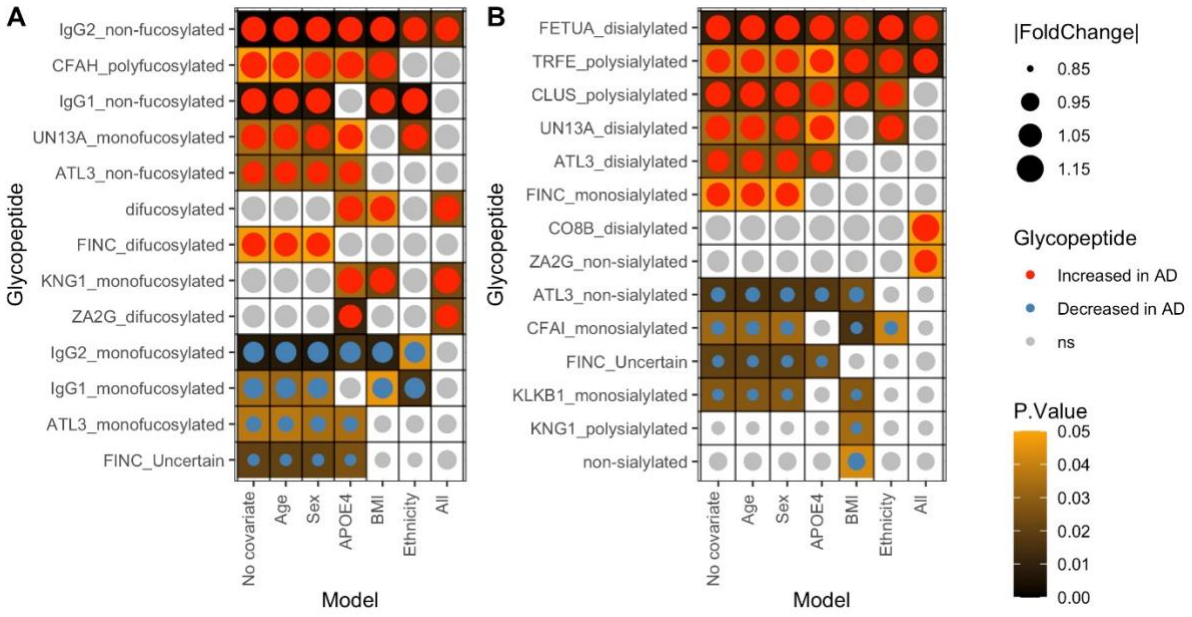
Supplemental Figure 3-S1 An overview of the glycoproteomic workflow.



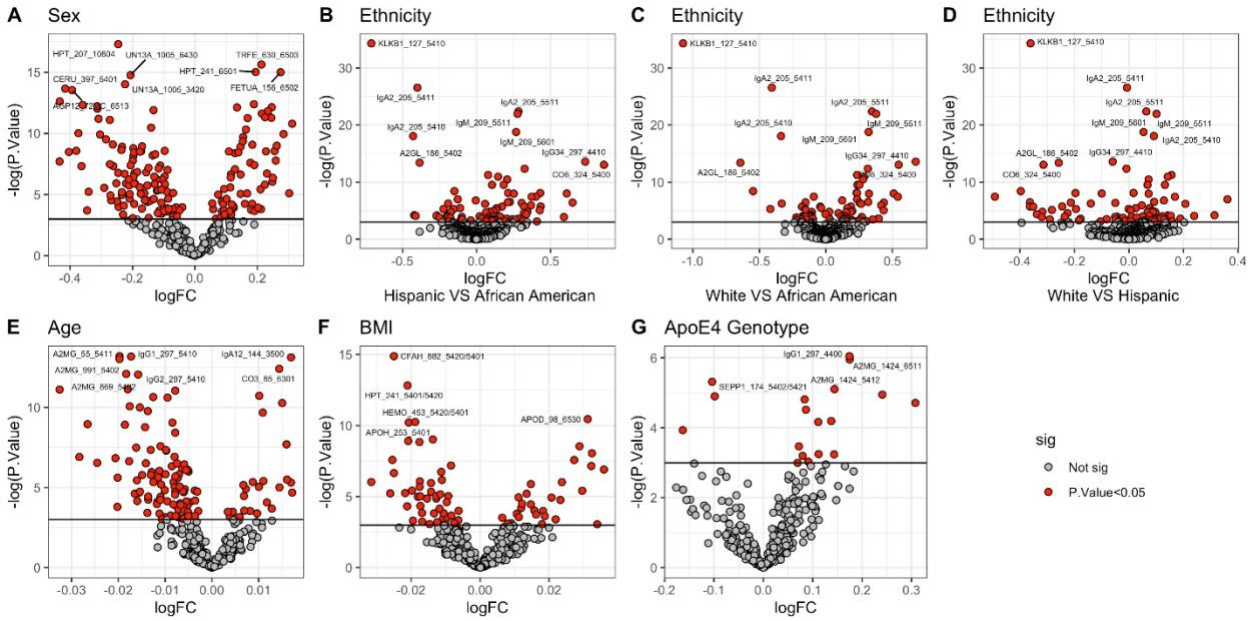
Supplemental Figure 3-S2 Glycopeptide sample outliers identified from hierarchical clustering.



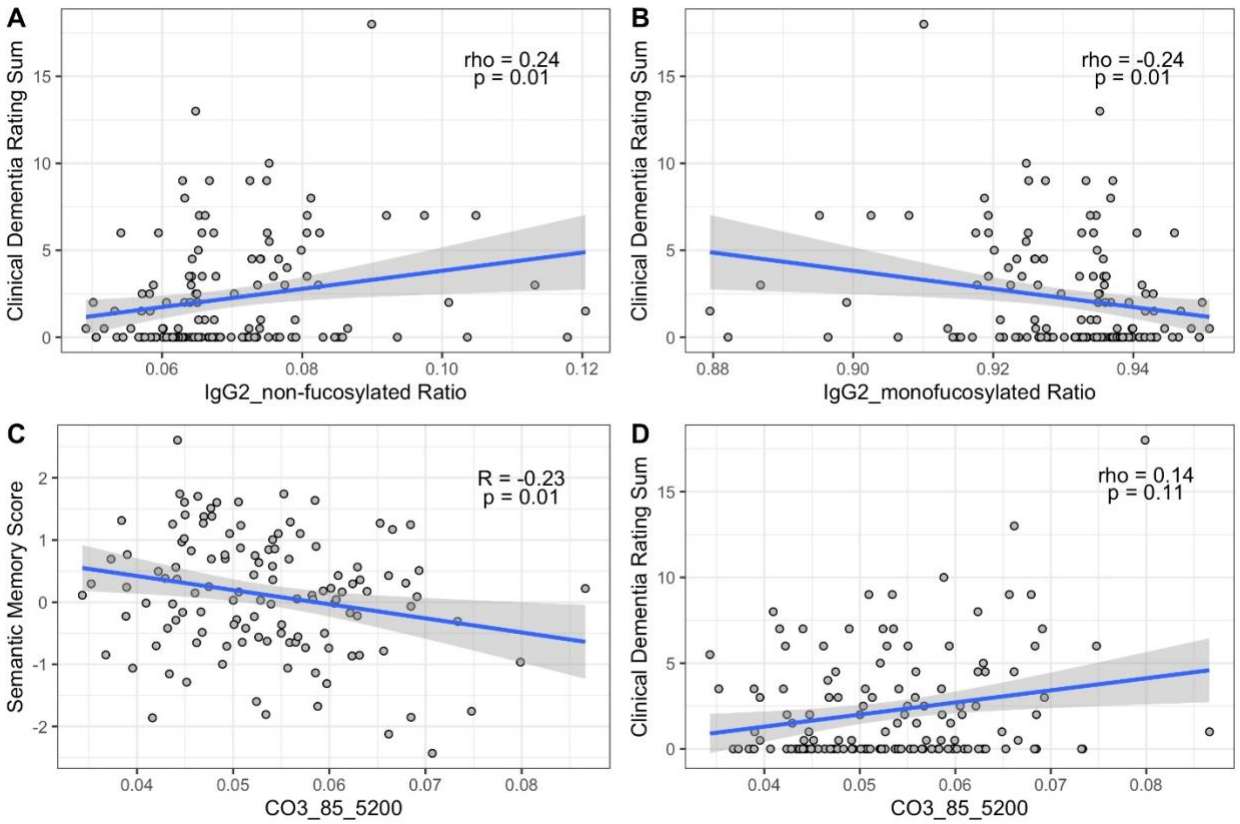
Supplemental Figure 3-S3 The distribution of the CVs of glycopeptides. The red vertical lines showed the cutoff at 0.2 (20%) and 0.3 (30%).



Supplemental Figure 3-S4 Differential analysis on the fucosylated and sialylated status of glycopeptides. **A-B**) Heatmaps presented the glycopeptides that differed in AD and control in at least one linear model ($p<0.05$) by their fucosylation status (**A**) and sialylation status (**B**). The dot size indicates the fold change of the glycopeptide in AD compared to control. Dot color indicates glycopeptides in the serum of AD patients that were down- (blue), upregulated (red), or not significantly changed (grey) compared to controls. The darkness of the background color shows the P-value.



Supplemental Figure 3-S5 The impact of potential confounders on glycopeptide abundance. Red dots showed glycopeptides impacted by each confounder ($p<0.05$), while grey dots were glycopeptides not impacted. **A**) Volcano plot showing the impact of sex on glycopeptides. Positive and negative log-fold changes indicate that the serum glycopeptide abundances in AD patients were higher or lower in females than males. The glycopeptides with -log(P-value) > 13 were labeled. **B-D**) Volcano plot showing the impact of ethnicity on glycopeptides with pairwise comparison. The glycopeptides with -log(P-value) > 13 were labeled. **E**) Volcano plot showing the impact of age on glycopeptides. The glycopeptides with -log(P-value) > 12 were labeled. **F**) Volcano plot showing the impact of BMI on glycopeptides. The glycopeptides with -log(P-value) > 10 were labeled. **G**) Volcano plot showing the impact of ApoE4 genotype on glycopeptides. Positive and negative log-fold changes indicate that the serum glycopeptide abundances in AD patients were higher or lower in ApoE4 positive individuals than ApoE4 negative ones. The glycopeptides with -log(P-value) > 5 were labeled.



Supplemental Figure 3-S6 Correlations between glycopeptides and cognitive function measures. A-B) Correlations between clinical dementia rating (CDR) sum and IgG2 fucosylation status. C-D) Correlations between CO3_85_5200 and semantic memory score and CDR sum. Protein abbreviations: Immunoglobulin G2 (IgG2) and Complement C3 (CO3).

REFERENCES

[1] Association A. 2021 ALZHEIMER ' S DISEASE FACTS AND FIGURES Race , Ethnicity and Alzheimer ' s in America. Alzheimers Dement 2021;13:1–104.

[2] Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's Dement 2018;14:535–62. https://doi.org/10.1016/j.jalz.2018.02.018.

[3] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's Dement 2011;7:263–9. https://doi.org/10.1016/j.jalz.2011.03.005.

[4] Molinuevo JL, Ayton S, Batrla R, Bednar MM, Bittner T, Cummings J, et al. Current state of Alzheimer's fluid biomarkers. Acta Neuropathol 2018. https://doi.org/10.1007/s00401-018-1932-x.

[5] Akasaka-Manya K, Manya H. The role of APP o-glycosylation in Alzheimer's disease. Biomolecules 2020;10:1–14. https://doi.org/10.3390/biom10111569.

[6] Schedin-Weiss S, Winblad B, Tjernberg LO. The role of protein glycosylation in Alzheimer disease. FEBS J 2014;281:46–62. https://doi.org/10.1111/febs.12590.

[7] Schedin-Weiss S, Gaunitz S, Sui P, Chen Q, Haslam SM, Blennow K, et al. Glycan biomarkers for Alzheimer disease correlate with T-tau and P-tau in cerebrospinal fluid in subjective cognitive impairment. FEBS J 2020;287:3221–34. https://doi.org/10.1111/febs.15197.

[8] Zhang Q, Ma C, Chin LS, Li L. Integrative glycoproteomics reveals protein n-glycosylation aberrations and glycoproteomic network alterations in Alzheimer's disease. Sci Adv 2020;6:1–19. https://doi.org/10.1126/sciadv.abc5802.

[9] Kirmiz C, Li B, An HJ, Clowers BH, Chew HK, Lam KS, et al. A serum glycomics approach to breast cancer biomarkers. Mol Cell Proteomics 2007;6:43–55. https://doi.org/10.1074/mcp.M600171-MCP200.

[10] Ruhaak LR, Kim K, Stroble C, Taylor SL, Hong Q, Miyamoto S, et al. Protein-Specific Differential Glycosylation of Immunoglobulins in Serum of Ovarian Cancer Patients. J Proteome Res 2016;15:1002–10. https://doi.org/10.1021/acs.jproteome.5b01071.

[11] Hong Q, Lebrilla CB, Miyamoto S, Ruhaak LR. Absolute quantitation of immunoglobulin G and its glycoforms using multiple reaction monitoring. Anal Chem 2013;85:8585–93. https://doi.org/10.1021/ac4009995.

[12] Hong Q, Ruhaak LR, Stroble C, Parker E, Huang J, Maverakis E, et al. A Method for Comprehensive Glycosite-Mapping and Direct Quantitation of Serum Glycoproteins. J Proteome Res 2015;14:5179–92. https://doi.org/10.1021/acs.jproteome.5b00756.

[13] Li Q, Kailemia MJ, Merleev AA, Xu G, Serie D, Danan LM, et al. Site-Specific Glycosylation Quantitation of 50 Serum Glycoproteins Enhanced by Predictive Glycopeptidomics for Improved Disease Biomarker Discovery. Anal Chem 2019;91:5433–45. https://doi.org/10.1021/acs.analchem.9b00776.

[14] Wu Z, Serie D, Xu G, Zou J. PB-Net: Automatic peak integration by sequential deep learning for multiple reaction monitoring. J Proteomics 2020;223:103820. https://doi.org/10.1016/j.jprot.2020.103820.

[15] Biermann MHC, Griffante G, Podolska MJ, Boeltz S, Stürmer J, Muñoz LE, et al. Sweet but dangerous - The role of immunoglobulin G glycosylation in autoimmunity and inflammation. Lupus 2016;25:934–42. https://doi.org/10.1177/0961203316640368.

[16] Plomp R, Ruhaak LR, Uh HW, Reiding KR, Selman M, Houwing-Duistermaat JJ, et al. Subclass-specific IgG glycosylation is associated with markers of inflammation and metabolic health. Sci Rep 2017;7:1–10. https://doi.org/10.1038/s41598-017-12495-0.

[17] Zhang X, Yuan H, Lyu J, Meng X, Tian Q, Li Y, et al. Association of dementia with immunoglobulin G N-glycans in a Chinese Han Population. Npj Aging Mech Dis 2021;7:1–12. https://doi.org/10.1038/s41514-021-00055-w.

[18] Maverakis E, Kim K, Shimoda M, Gershwin ME, Patel F, Wilken R, et al. Glycans in the immune system and The Altered Glycan Theory of Autoimmunity: A critical review. J Autoimmun 2015;57:1–13. https://doi.org/10.1016/j.jaut.2014.12.002.

[19] Krishnan S, Shimoda M, Sacchi R, Kailemia MJ, Luxardi G, Kaysen GA, et al. HDL glycoprotein composition and site-specific glycosylation differentiates between clinical groups and affects IL-6 secretion in lipopolysaccharide-stimulated monocytes. Sci Rep 2017;7:1–15. https://doi.org/10.1038/srep43728.

[20] Zhu C, Wong M, Li Q, Sawrey-Kubicek L, Beals E, Rhodes CH, et al. Site-Specific Glycoprofiles of HDL-Associated ApoE are Correlated with HDL Functional Capacity and Unaffected by Short-Term Diet. J Proteome Res 2019;18:3977–84. https://doi.org/10.1021/acs.jproteome.9b00450.

[21] Morris JC, Schindler SE, McCue LM, Moulder KL, Benzinger TLS, Cruchaga C, et al. Assessment of Racial Disparities in Biomarkers for Alzheimer Disease. JAMA Neurol 2019;76:264–73. https://doi.org/10.1001/jamaneurol.2018.4249.

[22] Hampel H, O'Bryant SE, Molinuevo JL, Zetterberg H, Masters CL, Lista S, et al. Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. Nat Rev Neurol 2018;14:639–52. https://doi.org/10.1038/s41582-018-0079-7.

[23] Tarawneh R. Biomarkers: Our Path Towards a Cure for Alzheimer Disease. Biomark Insights 2020;15. https://doi.org/10.1177/1177271920976367.

[24] Henriksen K, Bryant SEO, Hampel H, Trojanowski JQ, Montine TJ, Wyss-coray T, et al. The future of blood-based biomarkers for Alzheimer ' s disease. Alzheimer's Dement 2014;10:115–31.

Chapter 4: Quantitative Glycoproteomics of High-Density Lipoproteins

Xinyu Tang, [1] Maurice Wong, [2] Jennyfer Tena, [2] Chenghao Zhu,[1] Christopher Rhodes, [1] Qingwen Zhou, [2] Anita Vinjamuri, [2] Armin Oloumi, [2] Sucharita Boddu, [3] Guillaume Luxardi, [3] Emanual Maverakis, [3] Carlito B. Lebrilla, [2] and Angela M. Zivkovic [*1]

[1]Department of Nutrition, University of California, Davis, One Shields Ave., Davis, CA 9516, US.

[2]Department of Chemistry, University of California, Davis, One Shields Ave., Davis, CA 95616, US.

[3]Department of Dermatology, University of California, Davis School of Medicine, 3301 C Street, Suite 1400, Sacramento, CA 95816, USA

E-mail: amzivkovic@ucdavis.edu

Published manuscript with links to all supplemental files available at:

https://pubs.rsc.org/en/content/articlelanding/2022/ra/d2ra02294j

ABSTRACT

In this work, we developed a targeted glycoproteomic method to monitor the site-specific glycoprofiles and quantities of the most abundant HDL-associated proteins using Orbitrap LC-MS for (glyco)peptide target discovery and QqQ LC-MS for quantitative analysis. We conducted a pilot study using the workflow to determine whether HDL protein glycoprofiles are altered in healthy human participants in response to dietary glycan supplementation.

INTRODUCTION

High-density lipoproteins (HDL) range in size from 8-12 nm in diameter and have a density between 1.063-1.21 g/mL[1]. HDL are most known for their critical role in regulating cholesterol concentrations, but they perform other protective functions, including immunomodulatory, antioxidant, and antiproteolytic, among other functions[2]. The HDL proteome is complex and dynamic. HDL-associated proteins assist in stabilizing particle structure and solubilizing the lipid component, participate in signaling and interactions with lipoprotein receptors, and serve as cofactors for enzymes of lipid metabolism[1]. Apolipoproteins and other HDL-associated proteins such as Apo-E, ApoA-IV, ApoA-V, Apo-M, and serum amyloid A (SAA) are also involved in immune responses to infection and inflammation[3–5]. Many of these proteins are glycosylated with N- and O- glycans[6,7], and they typically exhibit glycan structural heterogeneity at each glycosite. Glycosylation can modulate HDL protein function, affect lipid- and protein-binding affinities, and is correlated with the functional capacity of HDL[8–10]. The effects of glycosylation of most HDL proteins are still incompletely understood. There is a growing interest in the quantification of proteins and site-specific glycoforms of HDL particles[11].

Previous work from our group has established the applicability and reliability of the multiple reaction monitoring (MRM) method in the quantitative analysis of peptides and glycopeptides from serum[12,13] and purified HDL[6,7]. We expanded upon the previous methods to craft a workflow tailor-made for HDL by employing LC-MS with an Orbitrap analyzer to discover and identify new HDL-associated tryptic (glyco) peptides. We also optimized the tryptic digestion procedure to ensure reproducibility and sensitivity toward target glycopeptides. In the previous method we monitored 8 proteins, including 23 glycopeptides[7].

The new method includes 339 transitions, spanning 47 peptides and 170 glycopeptides from 33 proteins. Absolute quantitation was achieved for the proteins ApoA-I, ApoC-I, Apo-D, Apo-E, and clusterin (Clus) by calibration with commercially available protein standards (**Table 4-S1**).

We then used the new panel to determine whether dietary glycan supplementation affects HDL glycoprofiles. Twenty-two healthy adult men/women (age 18-45) with a Body Mass Index (BMI) range of 18.5-25 were randomized to four treatment groups: placebo (n=4), N-acetylglucosamine (n=6), *Spirulina* (n=6), and galactose (n=6). Blood samples were collected before and after four weeks of supplementation. HDL was isolated using an optimized, validated method[14], and peptides and glycopeptides were quantified using the new workflow.

MATERIALS AND METHODS

Sample preparation for untargeted glycoproteomics analysis

For untargeted glycoproteomics analysis, blood plasma samples collected from a previously conducted study in 18-month old children in Ghana[15]. Purified HDL particles were isolated through a two-step HDL isolation method which isolates HDL particles first by density using sequential flotation ultracentrifugation followed by size exclusion chromatography, as described previously[14]. Briefly, 500 µL of plasma was underlaid under KBr solution at a density of 1.0060 g/mL to remove triglyceride-rich, low density (<1.006 g/mL) particles, including chylomicrons and VLDL, and submitted to ultracentrifugation in an Optima MAX-TL Ultracentrifuge with (Beckmann-Coulter) fixed angle rotor at 110,000 RPM and 14 ℃ for 30 minutes. After centrifugation, the supernatant was removed by aspiration, and the remaining fraction containing HDL, LDL, albumin, and plasma proteins was adjusted to a density of 1.210 g/mL with 1.340 g/mL KBr solution and underlaid under clean 1.210 g/mL density solution, then submitted to ultracentrifugation at 110,000 RPM and 14 ℃ for 3 hours and 30 minutes. The supernatant was removed by aspiration and dialyzed using an Amicon Ultra-4 50kDa centrifugal filter (Millipore) by centrifugation at 4,500 RPM for 8 minutes. A final volume of 250 µL was then transferred to an amber vial for FPLC analysis using a single Superdex 200 Increase 10/300 GL agarose-crosslinked column (GE Healthcare) on an AKTA P-920 FPLC (Amersham Biosciences). Four 1-mL fractions of HDL were pooled together and dialyzed to 100 µL, of which one 25 µL aliquot was used for glycoproteomic analysis.

For sample digestion, the HDL fractions were diluted in a total of 100 µL with 50 mM ammonium bicarbonate buffer at pH 7.5. All reagents used for sample preparation were freshly prepared in a buffer of 50 mM ammonium bicarbonate. Proteins were then denatured with 2 µL

of 550 mM dithiothreitol (Promega, Madison, WI) for 1 h at 65 ℃ and alkylated with 4 µL of 450 mM iodoacetamide (Sigma-Aldrich, St. Louis, MO) for 30 min at room temperature away from light. Proteins were digested with 2 µg sequencing grade trypsin (Promega) for 18 h at 37℃. Samples were purified through Bond Elut C18 Solid phase extraction (Agilent, Santa Clara, CA), dried in a vacuum concentrator and reconstituted in 50 µL LC-MS grade water.

Untargeted glycoproteomics analysis

Samples were run on an Agilent 1290 Infinity II High Performance Liquid Chromatography (HPLC) coupled to a Fusion Lumos MS/MS Orbitrap (Thermo Fisher Scientific). The HPLC was equipped with a 150 mm Agilent Zorbax Eclipse Plus C18 column with 1.8 µm particle size. Peptides and glycopeptides were eluted with a binary gradient of (A) 3% acetonitrile with 0.1% formic acid in water, and (B) 90% acetonitrile with 0.1% formic acid in water. The HPLC was set at a flow rate of 0.3 mL/min and programmed to ramp from 0% to 20% B in 20 min, 30% at 40 min, 44% at 47 min, and 100% at 48 min followed by a flushing and equilibration cycle. The electrospray ionization (ESI) voltage was set to 3500 V in the positive mode.

The Orbitrap was operated in positive mode with a precursor scan resolution of 60,000 and a range of 350-2000 m/z. Fragmentation was accomplished by stepped High-energy Collisional Dissociation (HCD). The collision energy was set at 30% and stepped at 10%. Ions for fragmentation were filtered to include a precursor mass range of 700-2000 m/z and charge states between 2-6.

Peptides and glycopeptides were identified with Byonic software (Protein Metrics Inc). For glycopeptide identification, a database of protein sequences and a library of glycan compositions are required as inputs. The human proteome database was downloaded from

Uniprot.org. We used in-house libraries for N-glycan and O-glycan compositions. (Glyco)peptides were identified based on the accurate mass of the precursor ions with tolerance set at 10 ppm and by matching MS/MS fragmentation with theoretical MS/MS spectra generated from in silico digestion of the provided protein database.

Sample preparation for targeted quantitative glycoproteomics analysis

Tryptic digestion of the HDL samples was done in a 96-well format to facilitate batch processing. Samples were randomized prior to plating. All reagents were freshly prepared in a buffer of 50 mM ammonium bicarbonate. A sample volume of 10 µL purified HDL was used for tryptic digestion. After every 20 samples, 10 µL of commercially available human serum (Sigma-Aldrich) was also digested to serve as sample preparation controls. Protein standards (APOA1, APOC1 APOD, APOE, Clusterin; all from Sigma-Aldrich) were mixed in known amounts (250, 250, 125, 200, and 125 µg/mL respectively) and digested with the batch to serve as calibration standards. Serial dilution of the digested protein mixture provided the calibration curve for absolute quantitation of the 5 proteins. The digested mixture was diluted by factors of 160, 80, 40, 20, 16, 8, 4, 2, and 1 to obtain 9 calibration standards, from which calibration curves spanning 4 orders of magnitude were calculated.

After pipetting the samples, controls, and standards onto the 96-well plate, 10 µL of 100 mM dithiothreitol was added to each well to reduce the protein disulfide bonds. Protein denaturation was continued by heating in a water bath for 1 h at 65 °C. Samples were then alkylated with 5 µL of 360 mM iodoacetamide for 30 min at room temperature away from light. Excess iodoacetamide was quenched with 5 µL of 100 mM dithiothreitol. Proteins were digested with 10 µL of 200 µg/mL sequencing grade trypsin

for 18 h at 37°C. The digestion was stopped by acidifying the solution with 5 µL 10% (v/v) formic acid (Fluka). To account for batch variability and possible run-order effects, 5 µL of a 1 µg/mL synthetic peptide with sequence RPAIAINNPYVPR (Bionexus, Oakland, CA) was added as an internal standard. The final volume of the HDL digest is 50 µL, a 5-fold dilution of the purified HDL solution. The samples were injected into the LC-MS instrument without further cleanup.

Targeted glycoproteomics analysis using QqQ

(Glyco)peptides were quantified on an Agilent 1290 Infinity II LC system coupled to an Agilent 6495B Triple Quadrupole MS. Injection volumes were set at 5 µL for protein standards and HDL samples, and 1 µL for serum digests. A pooled sample of the digested serum was run after every 10 HDL samples to serve as quality control (QC). They were used to monitor the stability of the instrument and the reproducibility of the batch analysis.

The HPLC was equipped with a 150 mm Agilent Zorbax Eclipse Plus C18 column with 1.8 µm particle size. A C18 column guard was used to protect the column from buildup of lipids and other hydrophobic substances in the sample. A binary gradient of (A) 3% acetonitrile with 0.1% formic acid in water, and (B) 90% acetonitrile with 0.1% formic acid in water was set at a flow rate of 0.5 mL/min. The HPLC pump parameters were programmed to ramp from 0% to 20% B in 20 min, 30% at 40 min, 44% at 47 min, and 100% at 48 min followed by 12 min column flushing cycle with 100% B and 7 min equilibration at 100% A. The Electrospray Ionization (ESI) voltage was set to 3500 V in the positive mode.

A transition list for target analytes was created by combining previously reported transitions[6,12] with new transitions selected from untargeted glycoproteomics analysis. The instrument was run on Dynamic Multiple Reaction Monitoring (DMRM) mode to minimize the number of transitions being monitored at each scan cycle. For peptides, at least 2 product ions were selected for monitoring. Quantitation was based on the area of the more abundant product ion while the others are for qualitative identification. Product ions for glycopeptides were based on diagnostic glycan fragments. Glycans yield characteristic oxonium ions after Collision Induced Dissociation (CID) with mass-to-charge ratios (m/z) of 204.08, 274.09, and 366.14 for N-acetylhexosamine (HexNAc), N-acetylneuraminic acid (NeuAc) with loss of H2O, and hexose + HexNAc (Hex1HexNAc1) respectively.

Participants and Study Design of Dietary Supplementation Pilot Study

The study was a four-week, double-blinded, randomized, placebo-controlled study. Twenty-two healthy adults (eleven males and ten females, age 18-45) with a Body Mass Index (BMI) range of 18.5-25 were recruited to the study and if they passed screening for inclusion and exclusion criteria were consented and randomized to four treatment groups: placebo (4), N-acetylglucosamine (6), *Spirulina* (6), and galactose (6). Exclusion criteria were prescription medications, supplements, smoking, drinking, special diet, intense exercise pattern, special diet features, weight changes, medical conditions, chronic disease, infection, autoimmune conditions, glycosylation-related diseases, low blood hemoglobin, and severe food or drug allergic reactions. Women were further excluded if they had a pregnancy within the last year or used hormonal contraception. Women were also excluded if they were perimenopausal or postmenopausal. The participants were

randomized to take a dietary supplement daily for 4 weeks, either placebo, 2.8 grams N-AG, 3000 milligrams *Spirulina*, or 25 grams galactose. The placebo contained one teaspoon cellulose powder. These doses were determined by a primary in vitro study which showed effects on cell surface glycosylation at these doses of monosaccharide in the culture medium. All participants came in to the study center for a morning blood draw at baseline and after four weeks. Blood samples were processed within 1 hour of the blood draw and plasma was aliquoted and stored immediately at -80C until analysis. HDL particles were isolated from plasma as described above. The study was approved by the University of California Davis Institutional Review board and registered at clinicaltrials.gov (NCT05040204).

Statistics

Protein concentrations of APOC1, APOA1, APOD, APOE, and CLUS were calculated based on the calibration curves. The relative abundance of the rest of the proteins was determined from the ion counts of the indicator peptide for each protein, as described previously[13]. The relative abundance of each glycopeptide was calculated as the ratio of the glycopeptide ion abundance to the ion abundance of the indicator peptide for the parent protein, as described previously[13]. The coefficient of variation (CV) was calculated for all (glyco)peptides based on pooled samples.

All statistical analyses were conducted in R version 4.1.0. Glycopeptides undetected in >= 5 samples or with CVs >= 30% across all samples were excluded from further analysis. Any remaining unobserved values were imputed as the minimum observed values of the specific glycopeptide. The fucosylated and sialylated peptide fractions were calculated as the sum of relative glycopeptide abundance of non-, mono-,

di-, or poly-glycosylated peptides relative to that of the total peptides to evaluate the change in overall fucosylation and sialylation of each protein.

The effects of placebo/supplements on HDL-related (glyco)peptides and their fucosylation/sialylation status were quantified as the fold changes of relative abundance/fractions comparing post-treatment to baseline. The differences in fold changes of abundances between the four groups were analyzed using a one-way ANOVA test. A pair-wise differential abundance test was performed to identify (glyco)peptides differentially altered between placebo and each supplement using a linear model with the limma package[16]. The resulting p-values were adjusted with multiple testing corrections using the Benjamini-Hochberg method. The fold change data were log2 transformed before linear model analysis and the Shapiro-Wilk test was performed to confirm the normality. The same data transformations and linear model were used for analysis of APOC1, APOA1, APOD, APOE, and CLUS using the standard curve derived concentrations, but no multiple testing corrections were applied.

RESULTS AND DISCUSSION

Identification of (glyco)peptides on HDL

Target discovery of HDL-associated proteins was accomplished through bottom-up glycoproteomics using Orbitrap MS on a total of 80 samples from young children[17]. Proteomics of HDL and other nanoparticles have been performed previously[5,11,18–20]. We considered both N- and O- linked glycosylation in searching for potential glycopeptides. Site-specific identification of O-linked glycopeptides is more challenging because tryptic digestion often yields peptides with multiple possible O-glycosylation sites, compounding the microheterogeneity at each site. Several issues can arise: ambiguity in assigning glycan compositions and positions of glycosylation in a given peptide and the multiplicity of glycoforms which further divides the glycopeptide signal. We indicated the ambiguity in position and/or glycan composition when appropriate. In **Figure 4-1** we show some examples of peptide and glycopeptide identifications from MS/MS spectra.

For each HDL protein, we chose the top-ranked peptides with the highest abundances and good sequence coverage for possible MRM-based detection. Some of the MRM transitions were obtained from previous work[6,9,12] and validated in the current workflow. Here we describe some of the proteins and their glycoforms that were analyzed in the MRM method.

ApoA-I is the predominant protein constituent in HDL and an important structural component[1]. It is known to play a role in cholesterol efflux and to act as a cofactor for lecithin cholesterol acyltransferase (LCAT)[1]. Transitions from 2 ApoA-I peptides were included; absolute quantitation was based on the peptide with sequence LAEYHAK,

which showed better linearity as seen in **Figure 4-2**. ApoA-II, ApoA-IV， and ApoA-V are minor protein constituents that can affect lipoprotein metabolism. Three possible O-glycosylation sites in ApoA-II featuring sialylated Core 1 glycans were found. ApoC-I, ApoC-II, ApoC-III, and ApoC-IV have relatively low molecular weights of just around 10-15 kDa. ApoC-I is an inhibitor of cholesteryl ester transfer protein (CETP) and reduces the esterification of free fatty acids. ApoC-II is an activator of lipoprotein lipase[21]. ApoC-III is the most abundant C apolipoprotein in human plasma and is O-glycosylated at the threonine residue in position 94 (position 74 after signal peptide cleavage). In our transition list, we included 10 different glycan structures on this site, as well as some variants in the peptide sequence.

Apo-D is closely associated with LCAT and can act as a multi-functional transporter. In our study, we included two N-glycosylation sites in positions 65 and 98, featuring mostly complex-type glycoforms.

Apo-E has been heavily studied because of its links to Alzheimer's disease and cardiovascular disease [22,23]. In this sample set, two O-glycosylation sites were found at positions 215 and 307/308, both with sialylated Core 1 structures. Positions 307 and 308 feature a threonine and a serine, respectively; the exact site of attachment could not be disambiguated based on MS/MS data. Clusterin (Clus, also known as Apo-J) is a ubiquitous glycoprotein that acts as a versatile chaperone. Its expression is increased in Alzheimer's disease, and it is involved in the clearance of amyloid-β peptides[24]. Different kinds of tumors have also been found to overexpress the protein or exhibit aberrant Clus glycosylation[25]. Changes in the glycosylation of Clus have yielded

possible biomarkers for Alzheimer's disease[24] and breast cancer[25]. In this study, N-glycosylation with complex-type structures was monitored at positions 86, 291, and 374.

Apo-M binds a variety of lipids and is involved in lipid transport. We found one N-glycosylation site at position 135, for which 8 different glycan compositions were observed. SAA proteins are major acute-phase reactants that rapidly increase in concentration during acute inflammation. SAA proteins are polymorphic, and they can present various isoforms[26]. In this study, we included peptides from three variants: SAA1, SAA2, and SAA4. Seven N-Glycopeptides from a variant of SAA4 (Uniprot Accession Code B2R5G8) that has an N-glycosylated site in position 94 were also monitored. This glycosite does not exist in the more common SAA4 variant. We further included several other minor protein constituents in HDL for a total of 34 unique proteins. The complete list of proteins and their corresponding Uniprot accession codes is provided in **Table 4-S2**.

Optimization of digestion protocol and method parameters

The HDL sample preparation and digestion workflow were optimized to minimize sample dilution while maintaining the efficacy of tryptic digestion and achieving adequate detection and good linearity for both peptides and glycopeptides. The optimized method featured a final dilution factor of 5 and a final protein concentration of around 0.5 µg/µL. This allowed us to inject 5 µL of the digested HDL sample to meet the target peptide level. We avoided cleanup steps such as solid-phase extraction to avoid further perturbations on the sample. The high lipid content of the HDL sample and the relatively low dilution factor necessitated additional conditioning and washing of the HPLC C18

column. A complete list of the MRM transitions, including the precursor and product ion m/z, retention times, and collision energies, are listed in **Table 4-S3**.

A synthetic peptide serving as an internal standard (ISTD) was added to each sample with a final concentration of 1 µg/mL to monitor the signal stability of the batch analysis. The ISTD signals from the 80 HDL samples had a coefficient of variation (CV) of 10.5%, demonstrating the stability of the instrument response throughout the batch analysis. The ISTD signal was also used to normalize all the peptide and glycopeptide signals for more consistent quantitation. Normalization was done by dividing the response of each analyte by that of the ISTD from the same sample run, then multiplying by the average ISTD response. The normalized response was used for all subsequent data processing.

Absolute quantification was determined for five proteins: ApoA-I, ApoC-I, Apo-D, Apo-E, and Clus. **Figure 4-2** shows the calibration curves for the five proteins. From the mass per volume concentration of the proteins, we also determined the number of protein units per volume of purified HDL. The molecular weight of the protein without its signal peptide was used for this calculation. Calibration curve plots based on molar concentrations are shown in **Figure 4-S1**. Because all signals were normalized to that of the ISTD, we compared the relative molar response factors based on the slope of the linear equations. In **Figure 4-S2,** we present the distribution of protein concentrations from 80 individuals in terms of both weights per volume and molar concentration. The median molar abundance of ApoA-I in HDL was 80 times greater than that of ApoC-I. The CVs of the molar abundances ranged between 34% to 43% for ApoA-I, ApoC-I, and

Apo-D. Among the five proteins, Apo-E displayed the most variation at 76%, while the CV of Clus was surprisingly low at just 4.0%.

Batch reproducibility

Pooled serum QC samples were run after every 10 sample injections to monitor the performance of the instrument. A total of 14 QC samples were run throughout the batch. For the five quantified proteins, their CVs were 4.4%, 3.2%, 4.4%, 5.5%, and 6.2% for the absolute quantification of ApoA-I, ApoC-I, Apo-D, Apo-E, and Clus respectively.

Among the 47 peptides monitored, 34 had CVs of less than 15%, while an additional 4 had CVs between 15-30%. Among the 9 peptides that have high CVs > 30%, 4 can be considered redundant because they were from proteins with better scoring peptides, while 4 more are from extremely low abundance proteins (Apo-F, Apo-(A), ApoC-IV, ApoA-V). The signal for the peptide from Serum paraoxonase/arylesterase 1 protein (PON1) showed higher variability and lower signal compared to previous results[12]. We believe this result was an unintended effect of the lower sample volume during tryptic digestion, leading to lower cleavage efficiency. PON1 is known to bind strongly to phospholipids and might not be fully dissociated and/or denatured in the less dilute digestion conditions that we employed. Of the 168 glycopeptides monitored, 69 had CVs of less than 30%, while a further 24 had CVs between 30-50%. About 60 of the glycopeptides were not present in appreciable amounts in the QC standards and had mean signals lower than 1000 ion counts, and so had high CVs. Heatmaps of the peptide and glycopeptide ISTD-normalized responses from all quality control runs are shown in **Figure 4-S3**. Overall, the method showed good reproducibility throughout the digestion

and instrument analysis for the target proteins; the limitation was mostly due to the naturally low abundance of certain analytes.

Relative glycoprotein quantitation

Relative quantitation can be obtained for the other 32 proteins for which there were no standards. The ISTD-normalized responses of the protein peptides are shown in **Figure 4-3**. The normalized responses can be compared relative to the same peptide across all samples; the variability across samples is demonstrated by the heatmap. However, we cannot directly compare different peptides to each other or infer their actual concentrations based solely on their normalized responses because their molar response factors can vary by several orders of magnitude.

Relative glycopeptide comparisons were performed by normalizing the glycopeptide responses to the peptide response from the same protein to ensure that the changes observed for the specific glycoforms would not be affected by changes in protein abundance. In **Figure 4-S4,** we illustrate the variation in glycan expression among the 80 samples.

To our knowledge, this is the first method to comprehensively quantify both the proteins and their glycosylation alterations of isolated HDL particles. As a proof of principle, we applied the method to determine compositional alterations of the proteins and their glycoprofiles in HDL isolated from participants before vs. after 4 weeks of supplementation with different monosaccharides.

Effects of dietary supplementations on HDL proteins and glycoprofiles

The baseline characteristics of the participants in each experimental arm are summarized in **Table 4-S4**. The age, BMI, systolic, and diastolic blood pressure of subjects were not significantly different among experimental arms (ANOVA P value >0.05 for all characteristics). The significantly altered glycans are shown in **Figure 4-4**. The mono-fucosylated fraction of Alpha-1 Antitrypsin (A1AT) differed across treatment groups and significantly decreased after *Spirulina* treatment compared to placebo (log2FC = 0.68±0.77 vs. -0.91±0.70 after taking the placebo vs. *Spirulina* for four weeks, **Figure 4-4A**). The abundance of disialylated Clus increased after four weeks of galactose supplementation (log2FC = -0.23±0.88 vs. 1.09±1.12 after taking the placebo vs. galactose for four weeks**, Figure 4-4B**), while the sialylation fractions of other glycopeptides remain unchanged across all supplement groups. Mono-fucosylated ApoE decreased after galactose treatment compared to control (log2FC = 0.93±1.30 vs. -0.14±0.86 after taking the placebo vs. galactose for four weeks), and an N-glycan on A1AT, A1AT_107_6513, was decreased after the *Spirulina* supplement (log2FC = 0.61±1.30 vs. -0.98±1.35 after taking the placebo vs. *Spirulina* for four weeks), although the differences of both glycopeptides between treatment group means did not reach statistical significance (**Figure 4-4C-D**). **Table 4-S5** lists all glycopeptides with differential fold changes comparing supplements to placebo.

As expected, the relative quantities of the proteins themselves were not affected by any supplement treatment. These results show for the first time that dietary glycan composition can affect HDL protein glycoprofiles without altering HDL protein concentrations, including changes in important functional proteins (i.e., Clus, Apo-E,

A1AT) known to be involved in a number of disease conditions, including Alzheimer's disease, metabolic disorders, and other chronic inflammatory diseases[24,27].

CONCLUSIONS

We have demonstrated the applicability of the developed HDL sample preparation and MRM MS workflow for the quantitation of proteins and glycopeptides in large clinical batches. Further developments will aim at the absolute quantitation of more target proteins for which purified standards can be obtained, as well as better sensitivity for glycopeptides from a variety of age groups and disease conditions. This method is sensitive enough to detect subtle glycosylation composition changes on HDL even in a pilot study with 4-6 subjects in each group.

DECLARATIONS

Author Contributions

A.M.Z, C.B.L, and E.M. conceived the project, M.W., J.T., C.Z., C.R., Q.Z., A.V., A.O., S.B., G.L. conducted the research, X.T. and M.W. analysed and visualized the data, X.T., M.W., J.T., and A.Z. wrote the paper.

Conflicts of interest

There are no conflicts to declare.

ACKNOWLEDGEMENTS

This work was supported by the National Institute of Health (R01GM049077 and R01AG062240) and National Institute of Food and Agriculture (CA-D-NUT-2242-H).

TABLES AND FIGURES

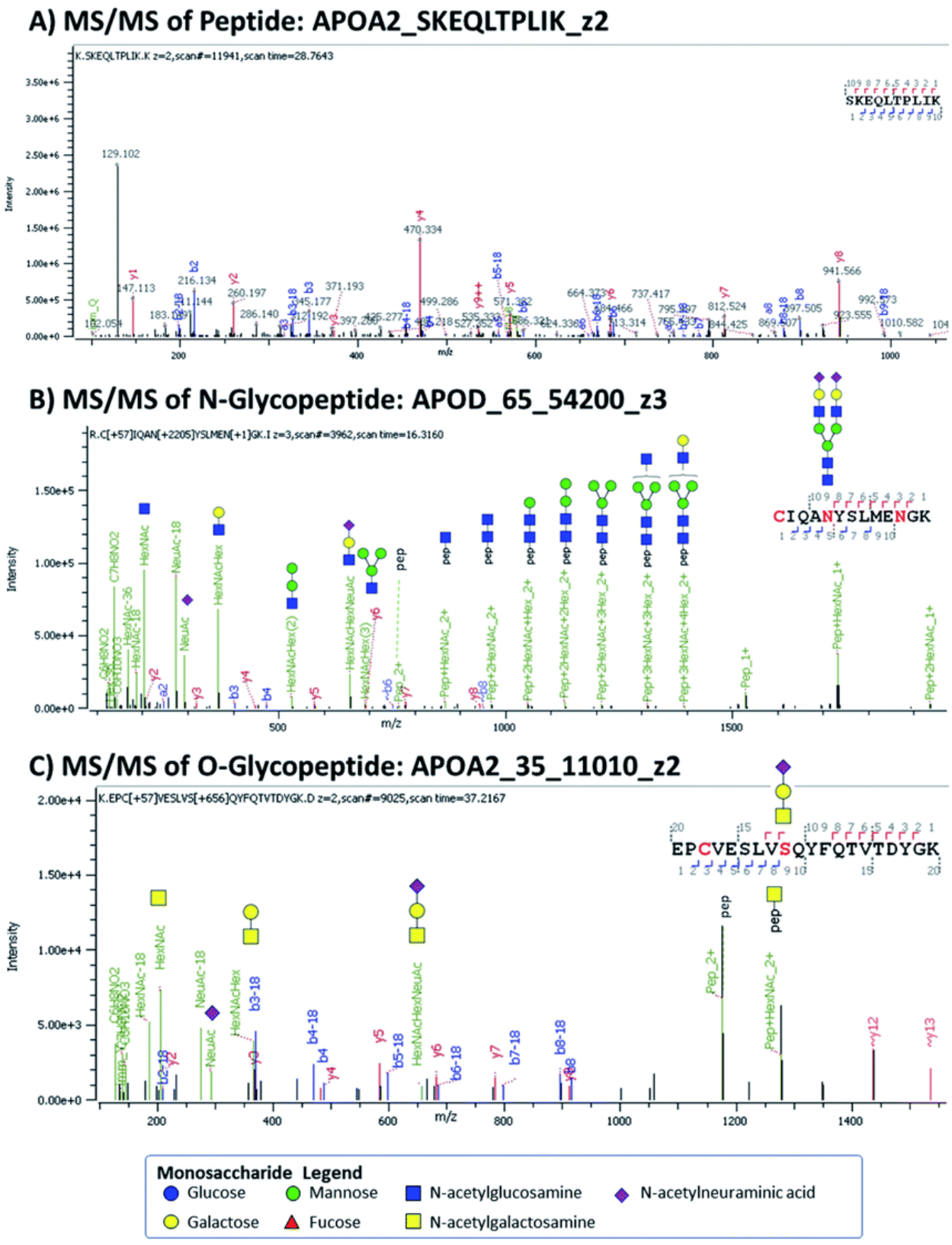
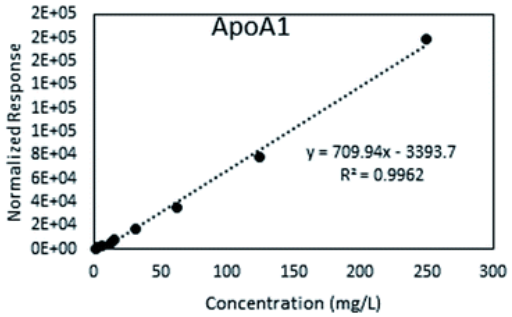
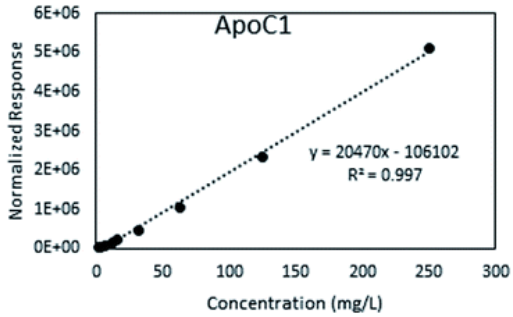
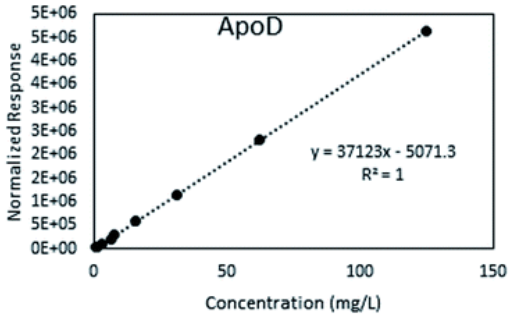
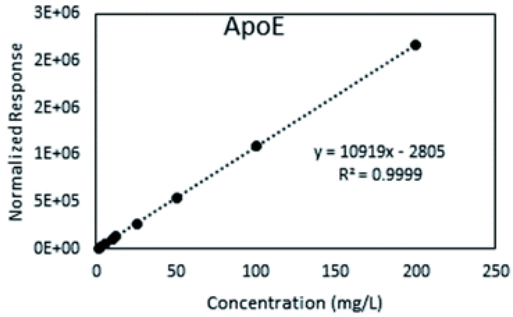


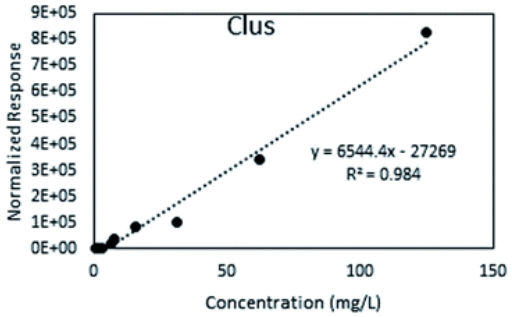
Figure 4-1 MS/MS spectra of (**A**) APOA2 peptide, (**B**) an N-glycopeptide from ApoD, and (**C**) an O-glycopeptide from ApoA2. The peptides are labeled as Protein_Sequence_ChargeState, while the glycopeptides are labeled as Protein_Position_GlycanComposition _ChargeState. Glycan compositions are written as 4- or 5-digit numbers denoting the number of hexose, HexNAc, fucose, NeuAc, and sulfate respectively.











Protein	Compound Name	Linear Equation	R^2
ApoA1	pep-APOA1_LAEYHAK	y = 710x - 3394	0.9962
ApoC1	pep-APOC1-QSELSAK	y = 20470x - 106102	0.9970
ApoD	QuantPep-APOD_VLNQELR	y = 37123x - 5071	1.0000
ApoE	pep-APOE_LGPLVEQGR_z2	y = 10919x - 2805	0.9999
Clus	QuantPep-CLUS_ASSIIDELFQDR	y = 6544x - 27269	0.9840

Figure 4-2 Calibration curves for Apo A-I, Apo C-I, Apo D, Apo E, and Clusterin/Apo J.

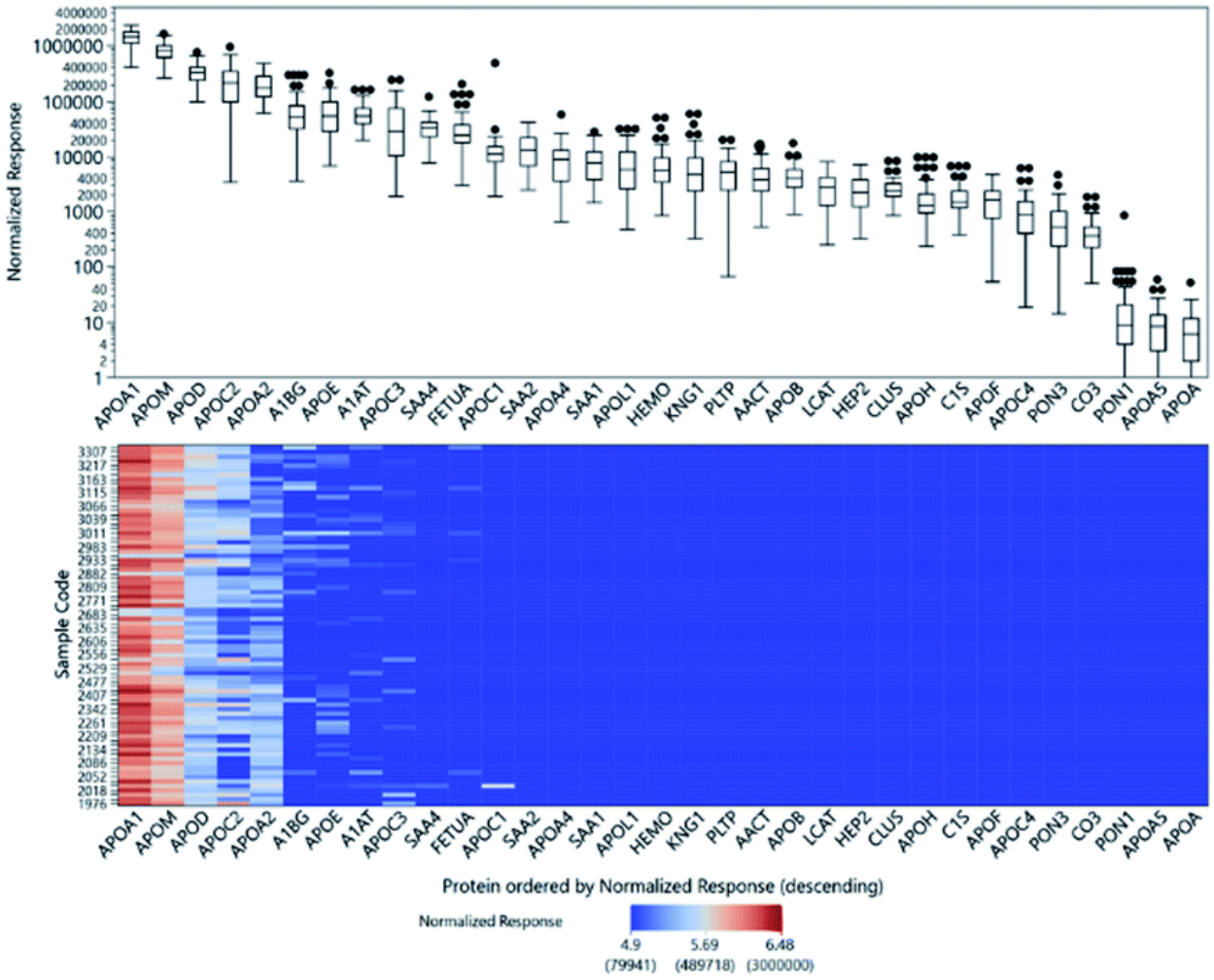


Figure 4-3 Top: Box plot of the ISTD-normalized responses of protein peptides showing the distribution of relative abundances from 80 individuals. The y-axis is plotted in the log 10 scale. **Bottom**: Heatmap of responses of protein peptides showing the ISTD-normalized response for each of the 80 samples. The color scaling is also done in the log 10 scale.

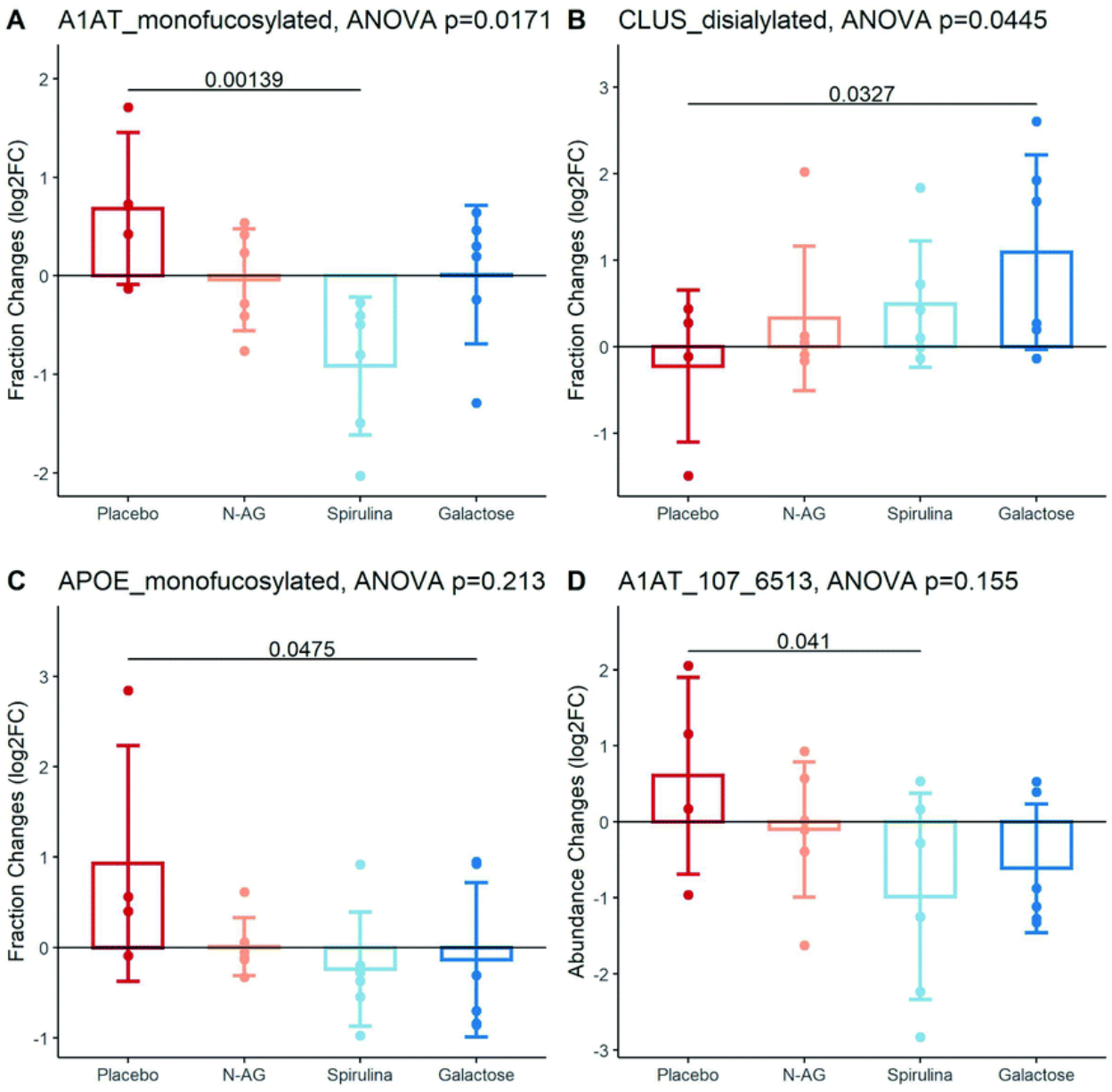


Figure 4-4 HDL glycan composition changes in response to supplements. (**A**) Fucosylation fraction changes on A1AT after treatment. (**B**) Sialylation fraction changes on CLUS after treatment. (**C**) Fucosylation fraction changes on APOE after treatment. (**D**) HDL glycan, A1AT_107_6513, changes after treatment. Unadjusted p-values were labeled.

Table 4-S1: The comparison between current HDL glycopeptide quantification method to previous methods

	J.Huang et al.	Current method
HDL proteins (n)	8	33
HDL peptides (n)	NA	47
HDL glycopeptides (n)	23	170
Absolute quantified proteins (n)	0	5

Table 4-S2: List of proteins measured in the optimized MRM method.

Abbrev	Common Name	Uniprot code
APOA	Apolipoprotein(A)	Q9UIR5
APOA2	Apolipoprotein A-II	P02652
APOA4	Apolipoprotein A-IV	P06727
APOA5	Apolipoprotein A-V	Q6Q788
APOC2	Apolipoprotein C-II	P02655
APOC4	Apolipoprotein C-IV	P55056
APOF	Apolipoprotein F	Q13790
APOL1	Apolipoprotein L1	O14791
HPTR	Haptoglobin-related protein	P00739
LCAT	Phosphatidylcholine-sterol acyltransferase	P04180
PLTP	Phospholipid transfer protein	P55058
A1AT	Alpha-1-antitrypsin	P01009
A1BG	Alpha-1B-glycoprotein	P04217
AACT	Alpha-1-antichymotrypsin	P01011
APOA1	Apolipoprotein A-I	P02647
APOB	Apolipoprotein B-100	P04114
APOC1	Apolipoprotein C-I	P02654
APOC3	Apolipoprotein C-III	P02656
APOD	Apolipoprotein D	P05090
APOE	Apolipoprotein E	P02649
APOH	Beta-2-glycoprotein 1	P02749
APOM	Apolipoprotein M	O95445
C1S	Complement C1s subcomponent	P09871
CLUS	Clusterin	P10909
CO3	Complement C3	P01024
FETUA	Alpha-2-HS-glycoprotein	P02765
HEMO	Hemopexin	P02790
HEP2	Heparin cofactor 2	P05546
KNG1	Kininogen-1	P01042
PON1	Serum paraoxonase/arylesterase 1	P27169
SAA	Serum amyloid A protein (SAA4 variant)	B2R5G8
SAA1	Serum amyloid A-1 protein	P0DJI8
SAA2	Serum amyloid A-2 protein	P0DJI9

Table 4-S3: MRM Transitions and (glyco)peptide details.

Protein Group	Compound	Precursor Ion	Product Ion (Quant)	Product Ion (Qual)	Ret Time (min)	Delta Ret Time	Collision Energy	Peptide Sequence	Charge State	Glycan Position[1]	Glycan Composition (Hex_HexNAc_Fuc_NeuAc)
ISTD	ISTD_RPAIAINNPYVPR	494.3	631.4	534.3	16.3	1.5	12	RPAIAINNPYVPR	3		
ISTD	ISTD_RPAIAINNPYVPR_z2	740.9	1324.8	631.4	16.3	1.5	15	RPAIAINNPYVPR	2		
ISTD	ISTD_RPAIAINNPYVPR_z4	371	534.3	631.4	16.3	1.5	10	RPAIAINNPYVPR	4		
A1AT	A1AT_AVLTIDEK	444.8	718.4	605.3	12.2	1.5	11	AVLTIDEK	2		
A1AT	A1AT_70_5402	1078.5	366.1	274.1	46.6	1.5	26	QLAHQSN^{70}STNIFFSPVSIATAFAMLSLGTK	5	70	5402
A1AT	A1AT_70_5412	1107.7	366.1		45.3	1.5	27	QLAHQSN^{70}STNIFFSPVSIATAFAMLSLGTK	5	70	5412
A1AT	A1AT_107_5401	1122.4	366.1		43.2	1.5	30	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	5401
A1AT	A1AT_107_5402	1180.6	366.1		43.2	1.5	29	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	5402
A1AT	A1AT_107_5411	1151.6	366.1		42.9	1.5	28	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	5411
A1AT	A1AT_107_5412	1209.8	366.1		43.2	1.5	30	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	5412
A1AT	A1AT_107_6503	1311.8	366.1		43.6	1.5	33	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	6503
A1AT	A1AT_107_6512	1282.9	366.1		43.5	1.5	30	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	6512
A1AT	A1AT_107_6513	1341	366.1		43.5	1.5	34	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	6513
A1AT	A1AT_107_6501/6520	1195.4	366.1	274.1	43.2	1.5	30	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	6501/6520
A1AT	A1AT_107_6502/6521	1253.6	366.1		43.1	1.5	30	ADTHDEILEGLNFN^{107}LTEIPEAQIHEGFQELLR	5	107	6502/6521
A1AT	A1AT_271_5401	1223.9	366.1		29.6	1.5	30	YLGN^{271}ATAIFFLPDEGK	3	271	5401
A1AT	A1AT_271_5402	991.2	366.1		29.6	1.5	24	YLGN^{271}ATAIFFLPDEGK	4	271	5402
A1AT	A1AT_271MC_5402	1149.9	366.1		42.9	1.5	28	YLGN^{271}ATAIFFLPDEGKLQHLENELTHDIITK	5	271	5402
A1AT	A1AT_271_5412	1027.7	366.1		29.6	1.5	25	YLGN^{271}ATAIFFLPDEGK	4	271	5412
A1AT	A1AT_271MC_5412	1179.1	366.1		42.9	1.5	29	YLGN^{271}ATAIFFLPDEGKLQHLENELTHDIITK	5	271	5412
A1AT	A1AT_271_55110_z3	1339.9	366.1	274.1	29.5	1.5	22	YLGN^{271}ATAIFFLPDEGK	3	271	5511
A1AT	A1AT_271_6503	1155	274.1	366.1	29.8	1.5	30	YLGN^{271}ATAIFFLPDEGK	4	271	6503
A1BG	A1BG_LLELTGPK	435.8	644.4	757.4	15.5	1.5	11	LLELTGPK	2		
A1BG	A1BG_179_5421/5402	1209.5	366.1		24.4	1.5	36	EGDHEFLEVPEAQEDVEATFPVHQPGN^{179}YSCSYR	5	179	5421/5402
AACT	AACT_ADLSGITGAR	480.8	661.4	404.2	11.8	1.5	13	ADLSGITGAR	2		
AACT	AACT_106_7603	1407.3	366.1		20.6	1.5	35	FN^{106}LTETSEAEIHQSFQHLLR	4	106	7603
AACT	AACT_106_7604	1480.1	366.1		24.1	1.5	35	FN^{106}LTETSEAEIHQSFQHLLR	4	106	7604
AACT	AACT_127_5401	1020.4	366.1		25.1	1.5	32	TLN^{127}QSSDELQLSMGNAMFVK	4	127	5401
AACT	AACT_127_8500	1120.2	366.1		25.3	1.5	34	TLN^{127}QSSDELQLSMGNAMFVK	4	127	8500
AACT	AACT_271_6502	1441.6	366.1		19.6	1.5	35	YTGN^{271}ASALFILPDQDK	3	271	6502
AACT	AACT_271_6512	1118.2	366.1		19.5	1.5	35	YTGN^{271}ASALFILPDQDK	4	271	6512
AACT	AACT_271_7602	1172.7	366.1		19.2	1.5	35	YTGN^{271}ASALFILPDQDK	4	271	7602
AACT	AACT_271_7603	1245.8	366.1		19.7	1.5	35	YTGN^{271}ASALFILPDQDK	4	271	7603
APOA	APOA_NPDAVAAPYCYTR_z2	749.3	398.2	859.4	20.8	1.5	22	NPDAVAAPYCYTR	2		
APOA1	APOA1_DLATVYVDVLK	618.3	736.4	936.5	26.5	1.5	17	DLATVYVDVLK	2		
APOA1	APOA1_LAEYHAK	416.2	647.3	718.4	3.9	1.5	10	LAEYHAK	2		
APOA2	APOA2_AGTELVNFLSYFVELGTQPATQ_z2	1193.1	416.2	815.4, 945.5	47	1.5	38	AGTELVNFLSYFVELGTQPATQ	2		
APOA2	APOA2_SKEQLTPLIK_z2	578.9	941.6	470.3, 571.4	13	1.5	16	SKEQLTPLIK	2		
APOA2	APOA2_35_11010_z2	1503.7	274.1		33.9	3	49	EPCVESLVS^{35}QYFQTVTDYGK	2	35	1101
APOA2	APOA2_35_11020_z3	1099.8	274.1		33.2	1.5	35	EPCVESLVS^{35}QYFQTVTDYGK	3	35	1102
APOA2	APOA2_88_11010_z2	1457.2	274.1		45.5	1.5	48	AGTELVNFLS^{88}YFVELGTQPAT	2	88/95	1101
APOA2	APOA2_95_11010_z2	1521.2	274.1		45.3	1.5	50	AGTELVNFLSYFVELGT^{95}QPATQ	2	88/95	1101
APOA4	APOA4_LAPLAEDVR_z2	492.3	274.2	589.3, 702.4	13.8	1.5	13	LAPLAEDVR	2		
APOA4	APOA4_LGEVNTYAGDLQK_z2	704.4	300.2	631.3, 1009.5	14.4	1.5	21	LGEVNTYAGDLQK	2		
APOA5	APOA5_AQLLGGVDEAWALLQGLQSR_z3	709.1	801.5	200.1, 560.3	43.3	1.5	21	AQLLGGVDEAWALLQGLQSR	3		
APOB	APOB_AHLDIAGSLEGHLR_z3	496.9	209.1	482.3, 611.3	17.6	1.5	13	AHLDIAGSLEGHLR	3		
APOB	APOB_GFEPTLEALFGK	654.8	975.6	488.5	32.2	1.5	14	GFEPTLEALFGK	2		
APOB	APOB_185_52000_z3	1149.2	366.1		23.1	1.5	18	QVLFLDTVYGN^{185}CSTHFTVK	3	185	5200
APOB	APOB_983_5401	1366.6	366.1		24.2	1.5	35	QVFPGLNYCTSGAYSN^{983}ASSTDSASYYPLTGDTR	4	983	5401
APOB	APOB_983_54020_z4	1439.6	366.1	274.1	24.6	1.5	24	QVFPGLNYCTSGAYSN^{983}ASSTDSASYYPLTGDTR	4	983	5402
APOB	APOB_3224_54020_z3	982.7	274.1	366.1	3	1.5	15	SYN^{3224}ETK	3	3224	5402
APOB	APOB_3411_5401	1174.2	366.1		10.4	1.5	30	FVEGSHN^{3411}STVSLTTK	3	3411	5401
APOC1	APOC1_QSELSAK	381.7	305.2	418.3	3.6	1.5	9	QSELSAK	2		
APOC2	APOC2_STAAMSTYTGIFTDQVLSVLK_z3	745.1	446.3	1002.6, 1149.7	42.2	1.5	22	STAAMSTYTGIFTDQVLSVLK	3		
APOC2	APOC2_TYLPAVDEK_z2	518.3	237.1	265.1, 658.3	13.4	1.5	14	TYLPAVDEK	2		
APOC3	APOC3_DALSSVQESQVAQQAR	858.9	502.3	887.5	13.6	1.5	26	DALSSVQESQVAQQAR	2		
APOC3	APOC3_GWVTDGFSSLK	598.8	244.1	854.4	21.4	1.5	17	GWVTDGFSSLK	2		
APOC3	APOC3_94_0300	916.1	204.1		28.4	1.5	22	FSEFWDLDPEVRPT^{94}SAVAA	3	94	0300
APOC3	APOC3_94_0310	975.4	204.1		28.5	1.5	24	FSEFWDLDPEVRPT^{94}SAVAA	3	94	0310
APOC3	APOC3_94A_11000_z2	1251.6	366.1		29.4	1.5	40	FSEFWDLDPEVRPT^{94}SAVAA	2	94	1100
APOC3	APOC3_94AMC_11000_z3	915.8	366.1		28.4	1.5	28	DKFSEFWDLDPEVRPT^{94}SAVAA	3	94	1100
APOC3	APOC3_94_1101	931.8	274.1		29.4	1.5	22	FSEFWDLDPEVRPT^{94}SAVAA	3	94	1101
APOC3	APOC3_94A_11010_z2	1397.1	274.1		29.4	1.5	45	FSEFWDLDPEVRPT^{94}SAVAA	2	94	1101
APOC3	APOC3_94Aoff_1101	908.1	274.1		29.4	1.5	22	FSEFWDLDPEVRPT^{94}SAVA	3	94	1101
APOC3	APOC3_94MC_1101	989.1	274.1		28.4	1.5	24	DKFSEFWDLDPEVRPT^{94}SAVA	3	94	1101
APOC3	APOC3_94_1102	1028.8	274.1		30	1.5	25	FSEFWDLDPEVRPT^{94}SAVAA	3	94	1102
APOC3	APOC3_94A_11020_z2	1542.7	274.1		30	1.5	51	FSEFWDLDPEVRPT^{94}SAVAA	2	94	1102
APOC3	APOC3_94Aoff_1102	1005.1	274.1		30	1.5	24	FSEFWDLDPEVRPT^{94}SAVA	3	94	1102
APOC3	APOC3_94MC_1102	1109.8	274.1		29.2	1.5	27	DKFSEFWDLDPEVRPT^{94}SAVAA	3	94	1102
APOC3	APOC3_94MC_11020_z3	1086.2	274.1		29	1.5	34	DKFSEFWDLDPEVRPT^{94}SAVAA	3	94	1102
APOC3	APOC3_94_1111	980.4	274.1		30.1	1.5	24	FSEFWDLDPEVRPT^{94}SAVAA	3	94	1111
APOC3	APOC3_94_1202	1096.5	274.1		29.9	1.5	27	FSEFWDLDPEVRPT^{94}SAVAA	3	94	1202
APOC3	APOC3_94AMC_12020_z3	1177.5	274.1		26.3	1.5	38	DKFSEFWDLDPEVRPT^{94}SAVAA	3	94	1202
APOC3	APOC3_94_1210	951.1	366.1		26.9	1.5	23	FSEFWDLDPEVRPT^{94}SAVAA	3	94	1210
APOC3	APOC3_94_1300	970.1	366.1		29.6	1.5	23	FSEFWDLDPEVRPT^{94}SAVAA	3	94	1300
APOC3	APOC3_94_2110	937.4	366.1		29.3	1.5	22	FSEFWDLDPEVRPT^{94}SAVAA	3	94	2110
APOC3	APOC3_94AMC_22100_z3	1086.2	366.1		25.4	1.5	34	DKFSEFWDLDPEVRPT^{94}SAVAA	3	94	2210
APOC3	APOC3_94A_ungly_z2	1069	1097.6		28.5	2	17	FSEFWDLDPEVRPT^{94}SAVAA	2	94	Unglycosylated
APOC4	APOC4_GFMQTYYDDHLR_z3	515.9	177.1	425.3, 818.4	17.4	1.5	14	GFMQTYYDDHLR	3		
APOD	APOD_VLNQELR	436.3	659.4	545.3	8.9	1.5	11	VLNQELR	2		
APOD	APOD_65_54010_z3	1148.1	366.1	274.1	17.4	1.5	18	CIQAN^{65}YSLMENGK	3	65	5401
APOD	APOD_65_54020_z3	1239.1	366.1	274.1	17	1.5	20	CIQAN^{65}YSLMENGK	3	65	5402

[1] Denotes the position in the protein sequence that includes the signal peptide

Protein Group	Compound	Precursor Ion	Product Ion (Quant)	Product Ion (Qual)	Ret Time (min)	Delta Ret Time	Collision Energy	Peptide Sequence	Charge State	Glycan Position[1]	Glycan Composition (Hex_HexNAc_Fuc_NeuAc)
APOD	APOD_65_54200_z3	1148.5	366.1		17.7	1.5	18	CIQAN^{65}YSLMENGK	3	65	5420
APOD	APOD_65_65010_z3	1269.5	366.1	274.1	17.4	2	20	CIQAN^{65}YSLMENGK	3	65	6501
APOD	APOD_65_65020_z3	1366.5	366.1	274.1	17.4	2	22	CIQAN^{65}YSLMENGK	3	65	6502
APOD	APOD_65_65020_z3_deamidated	1360.9	366.1	274.1	17.2	1.5	22	CIQAN^{65}YSLMENGK	3	65	6502
APOD	APOD_65_65030_z3	1463.6	366.1	274.1	16.5	1.5	24	CIQAN^{65}YSLMENGK	3	65	6503
APOD	APOD_65_65030_z3_deamidated	1457.9	366.1	274.1	20	1.5	24	CIQAN^{65}YSLMENGK	3	65	6503
APOD	APOD_65_65210_z3	1366.9	366.1	274.1	16.3	1.5	22	CIQAN^{65}YSLMENGK	3	65	6521
APOD	APOD_65_76020_z3	1488.6	366.1	274.1	18	1.5	24	CIQAN^{65}YSLMENGK	3	65	7602
APOD	APOD_65_76210_z3	1488.9	366.1	274.1	18	1.5	24	CIQAN^{65}YSLMENGK	3	65	7621
APOD	APOD_98_54000_z3	1292.9	366.1		17	1.5	21	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	5400
APOD	APOD_98_54010_z3	1389.9	366.1	274.1	17.6	1.5	23	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	5401
APOD	APOD_98_5402	1115.7	366.1	274.1	18	1.5	34	ADGTVNQIEGEATPVN^{98}LTEPAK	4	98	5402
APOD	APOD_98_54020_z3	1487	366.1	274.1	18	1.5	24	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	5402
APOD	APOD_98_5410	1341.6	366.1		16.8	1.5	35	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	5410
APOD	APOD_98_54100_z3	1341.6	366.1		16.8	1.5	22	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	5410
APOD	APOD_98_54110_z3	1438.6	366.1	274.1	17.5	1.5	23	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	5411
APOD	APOD_98_54120_z3	1535.7	366.1	274.1	17.8	1.5	25	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	5412
APOD	APOD_98_65010_z3	1511.6	366.1	274.1	17.7	1.5	25	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	6501
APOD	APOD_98_6510	1098	366.1		21.7	1.5	34	ADGTVNQIEGEATPVN^{98}LTEPAK	4	98	6510
APOD	APOD_98_65110_z3	1560.3	366.1	274.1	17.1	1.5	26	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	6511
APOD	APOD_98_65130_z3	1754.1	366.1	274.1	17.6	1.5	29	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	6513
APOD	APOD_98_6520	1134.5	366.1		20.7	1.5	35	ADGTVNQIEGEATPVN^{98}LTEPAK	4	98	6520
APOD	APOD_98_6530	1171	366.1		22.3	1.5	35	ADGTVNQIEGEATPVN^{98}LTEPAK	4	98	6530
APOD	APOD_98_65320_z3	1754.4	366.1	274.1	17.2	2	29	ADGTVNQIEGEATPVN^{98}LTEPAK	3	98	6532
APOD	APOD_98_7600	1152.7	366.1		22.1	1.5	35	ADGTVNQIEGEATPVN^{98}LTEPAK	4	98	7600
APOD	APOD_98_9800	1335.3	366.1		21.8	1.5	39	ADGTVNQIEGEATPVN^{98}LTEPAK	4	98	9800
APOE	APOE_AATVGSLAGQPLQER_z2	749.4	642.4	898.5, 1155.6	14.9	1.5	22	AATVGSLAGQPLQER	2		
APOE	APOE_LGPLVEQGR	484.8	588.3		12.2	1.5	13	LGPLVEQGR	2		
APOE	APOE_215_11010_z2	1077.5	274.1		18.1	1.5	34	AATVGS^{215}LAGQPLQER	2	215	1101
APOE	APOE_215_11020_z2	1223.1	274.1		18.9	1.5	39	AATVGS^{215}LAGQPLQER	2	215	1102
APOE	APOE_307/308_11010_z2	1139	274.1		11	1.5	36	VQAAVGT^{307}S^{308}AAPVPSDNH	2	307/308	1101
APOE	APOE_307/308_11010_z3	759.7	274.1		11	1.5	23	VQAAVGT^{307}S^{308}AAPVPSDNH	3	307/308	1101
APOF	APOF_SGVQQLIQYYQDQK_z2	849.4	972.4	613.3, 1085.5	21.8	1.5	26	SGVQQLIQYYQDQK	2		
APOF	APOF_269_11010_z2	930.4	274.1		6.8	1.5	29	DANIS^{269}QPETTK	2	269	1101
APOF	APOF_273/274_11010_z3	772.4	274.1	931.5	9.3	2	23	DANISQPET^{273}T^{274}KEGLR	3	273/274	1101
APOF	APOF_273/274_11020_z3	869.4	274.1	931.5	10.2	2	26	DANISQPET^{273}T^{274}KEGLR	3	273/274	1102
APOH	APOH_TCPKPDDLPFSTVVPLK_z3	638.7	357.3	743.5, 814.3	23.6	1.5	18	TCPKPDDLPFSTVVPLK	3		
APOL1	APOL1_LTDVAPVSFFLVLDVVYLVYESK_z2	1308.7	500.3	526.2	47.5	1.5	42	LTDVAPVSFFLVLDVVYLVYESK	2		
APOL1	APOL1_VTEPISAESGEQVER_z2	815.9	651.3	1091.5, 1301.6	11.7	1.5	25	VTEPISAESGEQVER	2		
APOM	APOM_AFLLTPR	409.2	599.4	486.3	17.7	1.5	10	AFLLTPR	2		
APOM	APOM_SLTSCLDSK	505.8	810.4	709.3	10.6	1.5	13	SLTSCLDSK	2		
APOM	APOM_135_43010_z3	1360.9	274.1		21.4	3	22	TELFSSSCPGGIMLN^{135}ETGQG YQR	3	135	4301
APOM	APOM_135_44010_z3	1428.6	274.1		21.4	3	23	TELFSSSCPGGIMLN^{135}ETGQG YQR	3	135	4401
APOM	APOM_135_53010_z3	1414.9	274.1		21.1	3	23	TELFSSSCPGGIMLN^{135}ETGQG YQR	3	135	5301
APOM	APOM_135_54010_z3	1482.6	366.1	274.1	20.5	3	24	TELFSSSCPGGIMLN^{135}ETGQG YQR	3	135	5401
APOM	APOM_135_54020_z3	1579.6	366.1	274.1	21.4	3	26	TELFSSSCPGGIMLN^{135}ETGQG YQR	3	135	5402
APOM	APOM_135_5421	1185.3	366.1		26.5	3	36	TELFSSSCPGGIMLN^{135}ETGQG YQR	4	135	5421
APOM	APOM_135_55110_z3	1598.7	366.1	274.1	21.5	3	26	TELFSSSCPGGIMLN^{135}ETGQG YQR	3	135	5511
APOM	APOM_135_8500	1199.7	366.1		26.9	3	36	TELFSSSCPGGIMLN^{135}ETGQG YQR	4	135	8500
APOM	APOM_135_ungly_z2	1131.1	1296.6	1620.8	21.5	3	36	TELFSSSCPGGIMLN^{135}ETGQG YQR	2	135	Unglycosylated
C1S	C1S_LLEVPEGR	456.8	686.4	458.2	13.2	1.5	12	LLEVPEGR	2		
C1S	C1S_174_5401	1089.6	366.1		25	1.5	30	NCGVN^{174}CSGDVFTALIGEIASPNYPKPYPENSR	5	174	5401
C1S	C1S_174_5402	1147.8	366.1		25.4	1.5	30	NCGVN^{174}CSGDVFTALIGEIASPNYPKPYPENSR	5	174	5402
CLUS	CLUS_ASSIIDELFQDR	697.4	922.4	565.3	29.9	1.5	20	ASSIIDELFQDR	2		
CLUS	CLUS_86_6503	1270.2	366.1	274.1	6.4	1.5	38	EDALN^{86}ETR	3	86	6503
CLUS	CLUS_291_5400	856.3	366.1		3.4	2	25	HN^{291}STGCLR	3	291	5400
CLUS	CLUS_291_54020_z3	1050.4	366.1	274.1	3.5	1.5	33	HN^{291}STGCLR	3	291	5402
CLUS	CLUS_291_6503	952.1	366.1		3.7	1.5	30	HN^{291}STGCLR	4	291	6503
CLUS	CLUS_291_5420/5401	953.4	366.1	274.1	3.4	2	24	HN^{291}STGCLR	3	291	5420/5401
CLUS	CLUS_291_5421/5402	788.3	366.1		3.5	1.5	26	HN^{291}STGCLR	4	291	5421/5402
CLUS	CLUS_374_6512	1101.5	366.1		18.9	1.5	34	LAN^{374}LTQGEDQYYLR	4	374	6512
CLUS	CLUS_374_5402/5421	973.2	366.1		17.6	1.5	31	LAN^{374}LTQGEDQYYLR	4	374	5402/5421
CLUS	CLUS_374_5420/5401	1200.2	366.1		18.7	1.5	36	LAN^{374}LTQGEDQYYLR	3	374	5420/5401
CLUS	CLUS_374_6520/6501	1064.7	366.1		17.8	1.5	33	LAN^{374}LTQGEDQYYLR	4	374	6520/6501
CO3	CO3_TIYTPGSTVLYR	685.9	892.5	1156.6	18.1	1.5	20	TIYTPGSTVLYR	2		
CO3	CO3_85_5200	1157.9	204.1		19.7	1.5	30	TVLTPATN^{85}HMGNVTFTIPANR	3	85	5200
CO3	CO3_85_6200	1211.9	366.1		19.6	1.5	30	TVLTPATN^{85}HMGNVTFTIPANR	3	85	6200
CO3	CO3_85_6301	1032.7	366.1		20.7	1.5	30	TVLTPATN^{85}HMGNVTFTIPANR	4	85	6301
CO3	CO3_85_7200	1265.9	366.1		19.5	1.5	30	TVLTPATN^{85}HMGNVTFTIPANR	3	85	7200
FETUA	FETUA_AHYDLR	387.7	288.2	209.1	5.7	1.5	12	AHYDLR	2		
FETUA	FETUA_156_5400	1132.2	366.1		18.7	1.5	30	VCQDCPLLAPLN^{156}DTR	3	156	5400
FETUA	FETUA_156_5401	1229.2	366.1		19.7	1.5	31	VCQDCPLLAPLN^{156}DTR	3	156	5401
FETUA	FETUA_156_6410	1234.9	366.1		19.5	1.5	30	VCQDCPLLAPLN^{156}DTR	3	156	6410
FETUA	FETUA_156_6411	1331.9	366.1		19	1.5	30	VCQDCPLLAPLN^{156}DTR	3	156	6411
FETUA	FETUA_156_6502	1086.2	366.1		19.9	1.5	27	VCQDCPLLAPLN^{156}DTR	4	156	6502
FETUA	FETUA_156_6503	1159	366.1		20	1.5	29	VCQDCPLLAPLN^{156}DTR	4	156	6503
FETUA	FETUA_156_6513	1195.5	366.1		20.7	1.5	30	VCQDCPLLAPLN^{156}DTR	4	156	6513
FETUA	FETUA_156_5402/5421	995.1	366.1		20	1.5	24	VCQDCPLLAPLN^{156}DTR	4	156	5402/5421
FETUA	FETUA_176_5401	1070.4	366.1		20.6	1.5	26	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	5401
FETUA	FETUA_176_6501	1161.7	366.1		20.7	1.5	29	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	6501
FETUA	FETUA_176_6502	1234.3	366.1		21.7	1.5	31	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	6502
FETUA	FETUA_176_6503	1307.1	366.1		22	1.5	33	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	6503
FETUA	FETUA_176_6512	1271	366.1		21.8	1.5	32	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	6512
FETUA	FETUA_176_6513	1343.8	366.1		22	1.5	34	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	6513
FETUA	FETUA_176_7600	1180.5	366.1		21.8	1.5	29	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	7600
FETUA	FETUA_176_5402/5421	1143	366.1		21.8	1.5	28	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	5402/5421
FETUA	FETUA_176_5412/5431	1179.7	366.1		21.6	1.5	29	AALAAFNAQNN^{176}GSNFQLEEISR	4	176	5412/5431
FETUA	FETUA_346_1101	891.4	274.1		17	1.5	21	T^{346}VVQPS^{351}VGAAAGPVVPPCPGR	3	346	1101
FETUA	FETUA_346_2200	916.1	366.1		17.1	1.5	22	T^{346}VVQPS^{351}VGAAAGPVVPPCPGR	3	346	2200
HEMO	HEMO_NFPSPVDAAFR	610.8	775.4	959.5	20.5	1.5	22	NFPSPVDAAFR	2		
HEMO	HEMO_187_6503	1067.2	366.1		16.4	2	33	SWPAVGN^{187}CSSALR	4	187	6503

Protein Group	Compound	Precursor Ion	Product Ion (Quant)	Product Ion (Qual)	Ret Time (min)	Delta Ret Time	Collision Energy	Peptide Sequence	Charge State	Glycan Position[1]	Glycan Composition (Hex_HexNAc_ Fuc_NeuAc)
HEMO	HEMO_187_5402/5421	903.1	366.1		16.2	1.5	29	SWPAVGN^{187}CSSALR	4	187	5402/5421
HEMO	HEMO_187_5412/5431	1252.5	366.1		16.2	2	37	SWPAVGN^{187}CSSALR	3	187	5412/5431
HEMO	HEMO_453_5402/5421	1315.2	366.1		22.1	1.5	39	ALPQPQN^{453}VTSLLGCTH	3	453	5402/5421
HEMO	HEMO_453_5420/5401	1217.9	366.1		21.4	2	37	ALPQPQN^{453}VTSLLGCTH	3	453	5420/5401
HEMO	HEMO_240/246_5402	1015.2	366.1		5.3	2	30	N^{240}GTGHGN^{246}STHHGPEYMR	4	240/246	5402
HEP2	HEP2_IAIDLFK	410.3	635.4	706.4	22.2	1.5	10	IAIDLFK	2		
HEP2	HEP2_49_5402	1229.9	366.1		19.6	2	37	N^{49}LSMPLLPADFHK	3	49	5402
HEP2	HEP2_49_5412	1278.5	366.1		20.3	2	38	N^{49}LSMPLLPADFHK	3	49	5412
HPTR	HPTR_126_54010_z4	1149	366.1	274.1	23.1	1.5	18	MVSHHN^{126}LTTGATLINEQWLLTTAK	4	126	5401
HPTR	HPTR_126_54020_z4	1222.1	274.1	366.1	23.9	1.5	20	MVSHHN^{126}LTTGATLINEQWLLTTAK	4	126	5402
KNG1	KNG1_YFIDFVAR	515.8	720.4	607.3	23.3	1.5	14	YFIDFVAR	2		
KNG1	KNG1_169_5402	1208	366.1		23.8	1.5	36	HGIQYFNN^{169}NTQHSSLFMLNEVK	4	169	5402
KNG1	KNG1_205_6503	1069.4	366.1		13.3	1.5	33	ITYSIVQTN^{205}CSK	4	205	6503
KNG1	KNG1_205_6512	1033.9	366.1		13.6	2	32	ITYSIVQTN^{205}CSK	4	205	6512
KNG1	KNG1_205_5402/5421	905.6	366.1		13.3	1.5	29	ITYSIVQTN^{205}CSK	4	205	5402/5421
LCAT	LCAT_LEPGQQEEYYR_z2	706.3	585.3	243.1	10.8	1.5	21	LEPGQQEEYYR	2		
LCAT	LCAT_PVILVPGCLGNQLEAK_z2	854.5	1186.6	522.4	24.5	1.5	26	PVILVPGCLGNQLEAK	2		
LCAT	LCAT_SSGLVSNAPGVQIR_z2	692.9	669.4	941.5, 1040.6	15.5	1.5	20	SSGLVSNAPGVQIR	2		
LCAT	LCAT_108_65030_z3	1171.1	274.1	366.1	5	1.5	19	VVYN^{108}R	3	108	6503
PLTP	PLTP_FLEQELETITIPDLR_z2	909	500.3	288.2	34.1	1.5	28	FLEQELETITIPDLR	2		
PLTP	PLTP_GAFFPLTER_z2	519.3	276.1	762.4	7.7	1.5	14	GAFFPLTER	2		
POLG	POLG_2529_47510_z3	1203.8	274.1	366.1	10.9	1.5	19	N^{2529}LSGK	3	2529	4751
PON1	PON1_YVYIAELLAHK	660.4	529.3	781.5	23.8	1.5	19	YVYIAELLAHK	2		
PON1	PON1_253_4301	910.4	366.1		17.1	1.5	29	HAN^{253}WTLTPLK	3	253	4301
PON1	PON1_253_43010_z3	910.4	366.1	274.1	19.6	1.5	14	HAN^{253}WTLTPLK	3	253	4301
PON1	PON1_253_54020_z3	1129.1	274.1	366.1	17.9	1.5	18	HAN^{253}WTLTPLK	3	253	5402
PON1	PON1_253_5402/5421	847.4	366.1		16.8	1.5	28	HAN^{253}WTLTPLK	4	253	5402/5421
PON1	PON1_324_5420	1057.7	366.1		25.5	1.5	33	VTQVYAEN^{324}GTVLQGSTVASVYK	4	324	5420
PON1	PON1_324_6501	1149.3	366.1		25.4	1.5	35	VTQVYAEN^{324}GTVLQGSTVASVYK	4	324	6501
PON1	PON1_324_6502	1221.5	366.1		24.1	1.5	37	VTQVYAEN^{324}GTVLQGSTVASVYK	4	324	6502
PON1	PON1_324_65030_z3	1725.7	366.1	274.1	24	3	29	VTQVYAEN^{324}GTVLQGSTVASVYK	3	324	6503
PON3	PON3_IFLMDLNEQNPR_z2	745.4	261.2	514.3, 985.5	42.2	1.5	22	IFLMDLNEQNPR	2		
PON3	PON3_LLNYNPEDPPGSEVLR_z2	907	854.5	618.3, 1195.6	20.6	1.5	28	LLNYNPEDPPGSEVLR	2		
PON3	PON3_29_75310_z3	1142.4	366.1	274.1	19.5	3	18	VN^{29}ASR	3	29	7531
PON3	PON3_29_76210_z3	1161.5	366.1	274.1	19.5	3	19	VN^{29}ASR	3	29	7621
PON3	PON3_29_76030/76220_z3	1258.2	366.1	274.1	16.5	4	20	VN^{29}ASR	3	29	76030/76220
SAA	SAA_94_43010_z3	1391.6	366.1	274.1	43.2	3	23	VYLQGLIDYYLFGN^{94}SSTVLEDSK	3	94	4301
SAA	SAA_94_44010_z3	1459.3	366.1	274.1	43.2	3	24	VYLQGLIDYYLFGN^{94}SSTVLEDSK	3	94	4401
SAA	SAA_94_53010_z3	1445.6	366.1	274.1	43.2	3	24	VYLQGLIDYYLFGN^{94}SSTVLEDSK	3	94	5301
SAA	SAA_94_54010_z3	1513.3	366.1	274.1	43.2	3	25	VYLQGLIDYYLFGN^{94}SSTVLEDSK	3	94	5401
SAA	SAA_94_54020_z3	1610.4	366.1	274.1	43.2	3	27	VYLQGLIDYYLFGN^{94}SSTVLEDSK	3	94	5402
SAA	SAA_94_54210_z3	1610.4	366.1	274.1	43.2	3	27	VYLQGLIDYYLFGN^{94}SSTVLEDSK	3	94	5421
SAA	SAA_94_55100_z3	1532.3	366.1		43.2	3	25	VYLQGLIDYYLFGN^{94}SSTVLEDSK	3	94	5510
SAA1	SAA1_FFGHGAEDSLADQAANEWGR_z3	726.7	732.3	418.2, 948.4	20.2	1.5	21	FFGHGAEDSLADQAANEWGR	3		
SAA1	SAA1_GPGGVWAAEAISDAR_z3	486.2	448.2	561.3	19.8	1.5	13	GPGGVWAAEAISDAR	3		
SAA2	SAA2_GPGGAWAAEVISNAR_z3	485.9	447.2	560.3	19.8	1.5	13	GPGGAWAAEVISNAR	3		
SAA/SAA4	SAA4_FRPDGLPK_z2	465.3	244.2	516.3, 658.4	10.3	1.5	12	FRPDGLPK	2		

Table 4-S4: Baseline characteristics of subjects in each experimental arm.

Characteristics	Placebo	N-AG	*Spirulina*	Galactose	P value
n	4	6	6	6	
Sex = Male/Female	2/2	3/3	3/3	3/3	
Age (yrs) (median[IQR])	26[24-29]	27[25-28]	34[26-39]	28[27-31]	0.37
Ethnicity = Asian/Biracial/Caucasian/Hispanic	2/0/2/0	3/1/2/0	3/0/2/1	3/0/3/0	
Height (median[IQR])	65[62-68]	70[66-71]	66[64-69]	68[67-69]	0.5
Weight (lbs) (median[IQR])	137[125-152]	143[125-154]	148[137-160]	152[139-161]	0.81
BMI (kg/m2) (median[IQR])	24[23-24]	20[20-22]	24[22-25]	24[23-24]	0.07
Systolic Blood Pressure (median[IQR])	108[102-113]	117[102-120]	104[98-116]	115[109-117]	0.68
Diastolic Blood Pressure (median[IQR])	66[64-70]	73[71-77]	72[66-83]	70[68-76]	0.76

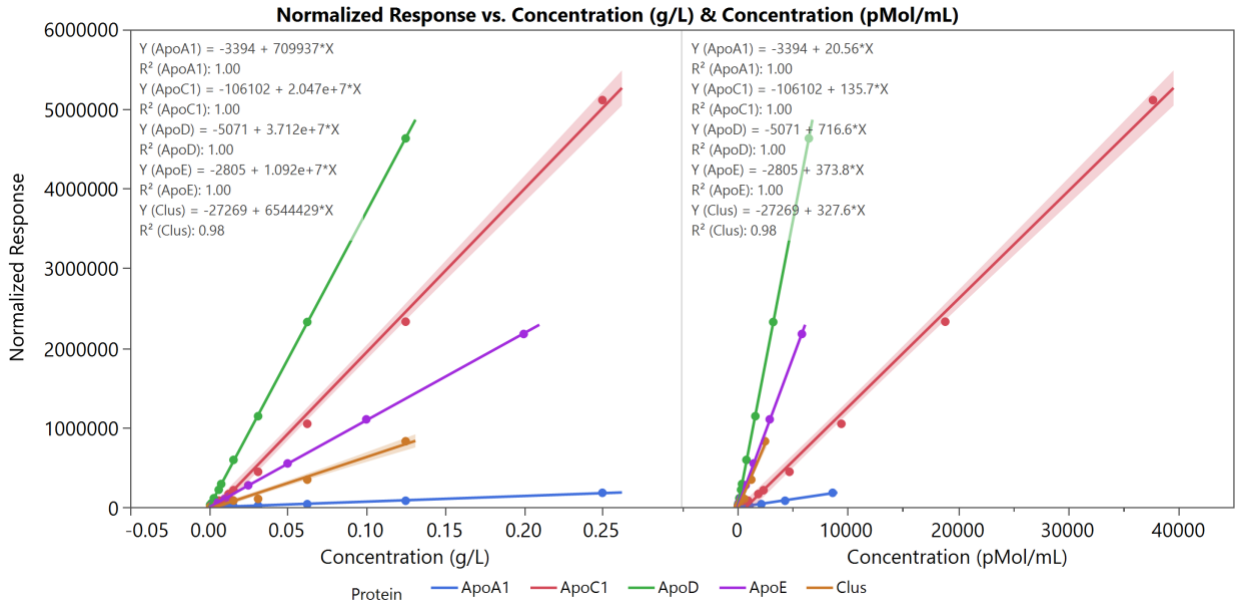
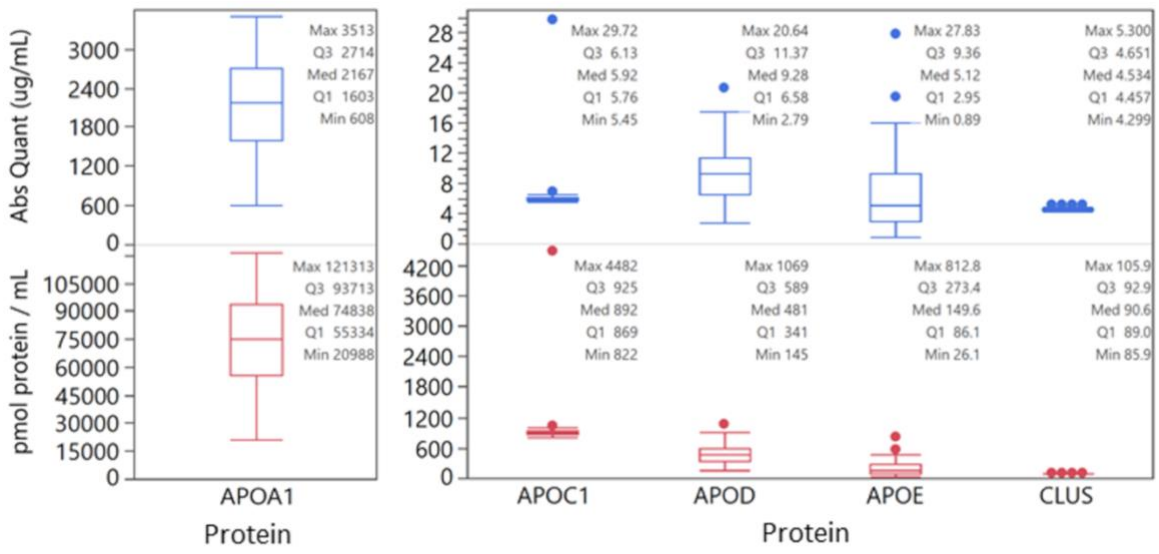


Figure 4-S1: Linear calibration curve plots for ApoA1, ApoC1, ApoD, ApoE, and Clus based on (left) mass per volume concentration and (right) molar concentration.



Protein	w/v Concentration (µg protein / mL purified HDL)		Molar Concentration (nM)		Coefficient of Variation (%)
	Median	Mean	Median	Mean	
APOA1	2167	2154	74838	74370	33.9
APOC1	5.92	6.24	892	941	42.9
APOD	9.28	9.27	481	480	38.1
APOE	5.12	6.45	150	189	76.3
CLUS	4.53	4.57	90.6	91.4	4.0

Figure 4-S2: Box plots of the absolute quantities of APOA1, APOC1, APOD, APOE, and Clus. The plots show the natural variability of the protein concentrations from 80 individuals.

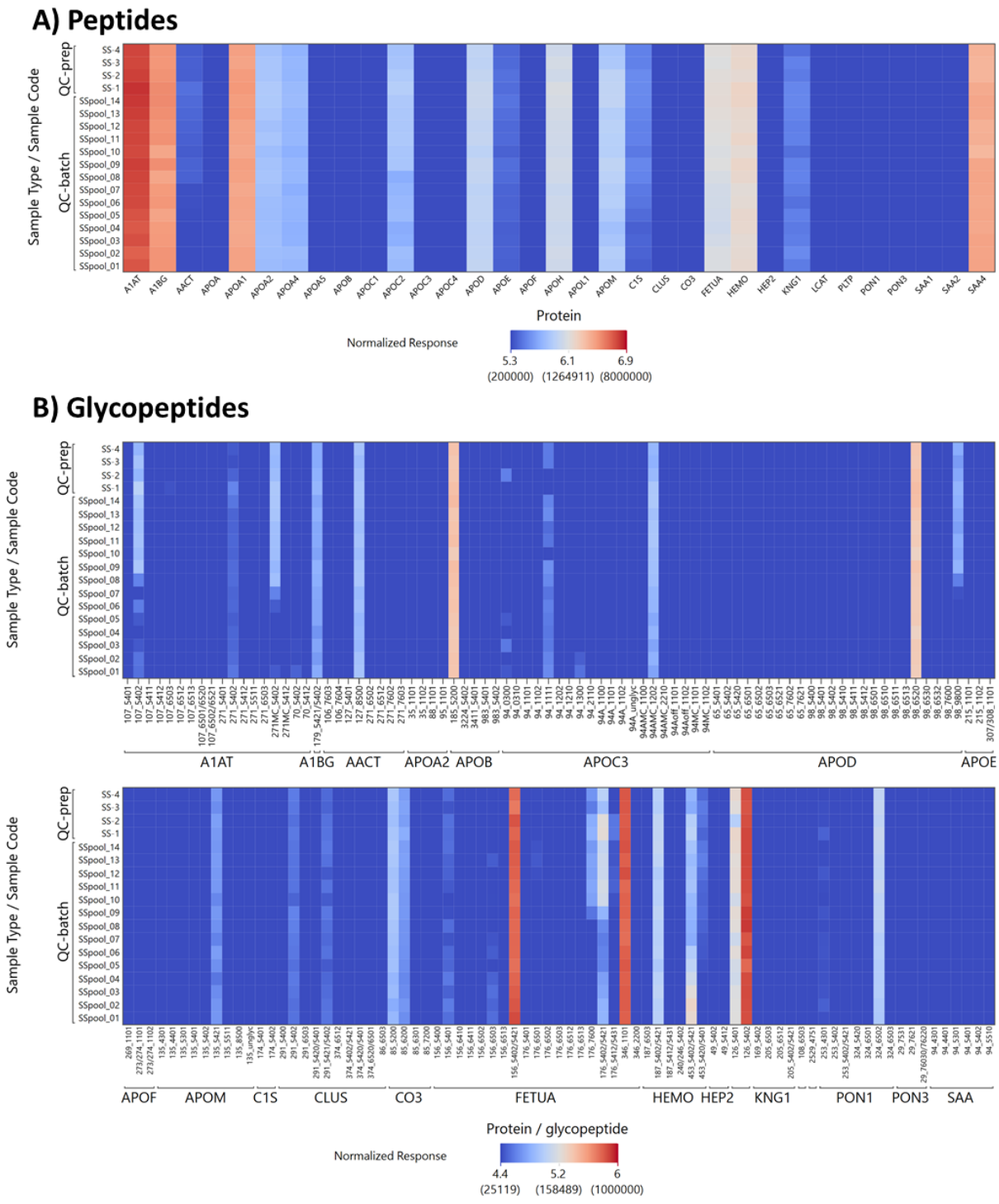


Figure 4-S3: Heatmaps of the ISTD-normalized responses for (A) peptides and (B) glycopeptides in quality control samples.

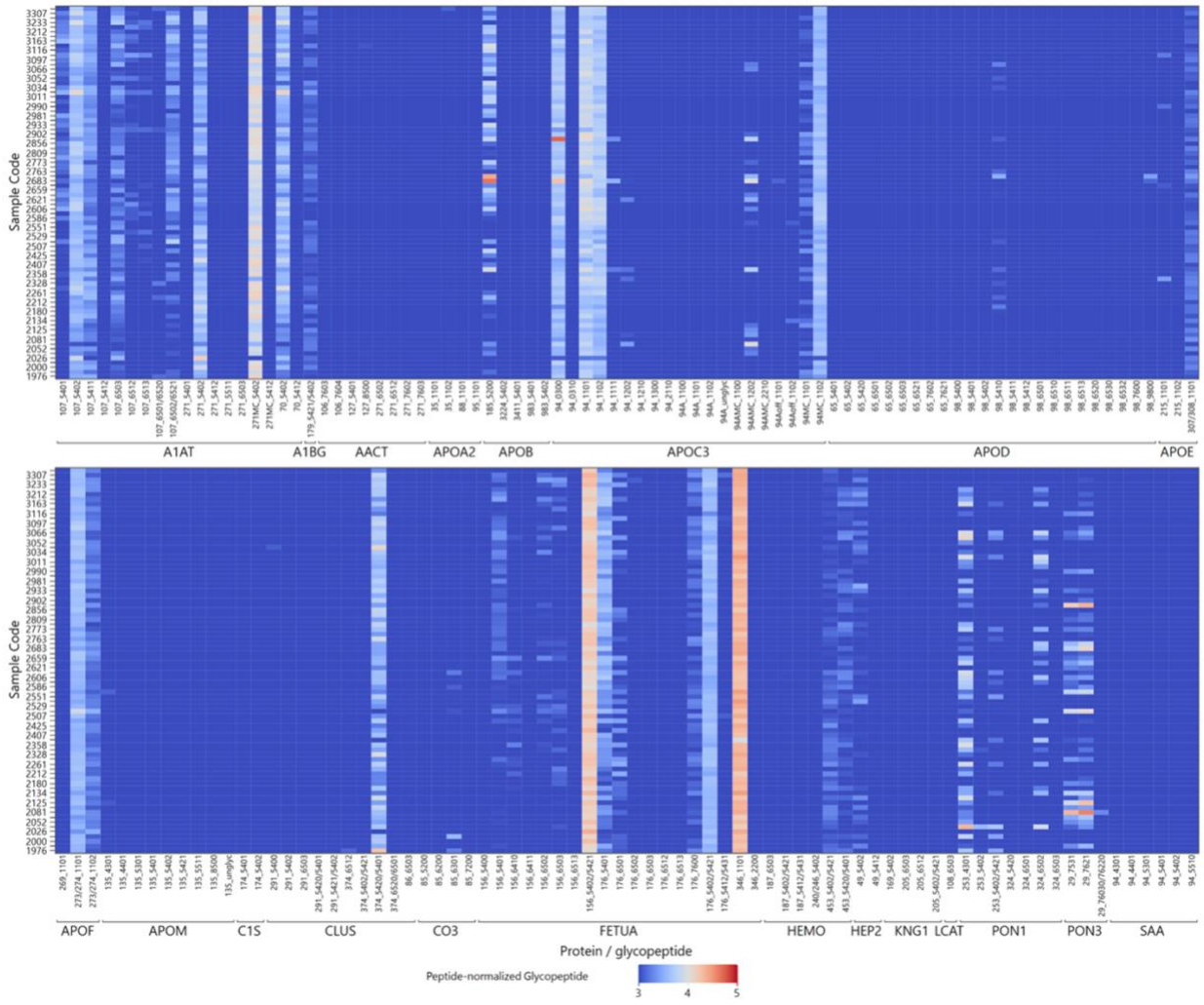


Figure 4-S4: Heatmap of glycopeptide expression showing the peptide-normalized response for 168 glycoforms in each of the 80 samples. The color scaling is in the log 10 scale.

REFERENCES

[1] Mahley RW, Innerarity TL, Rall SC, Weisgraber KH. Plasma lipoproteins: Apolipoprotein structure and function. J Lipid Res 1984;25:1277–94. https://doi.org/10.1016/s0022-2275(20)34443-6.

[2] Ansell BJ, Watson KE, Fogelman AM, Navab M, Fonarow GC. High-density lipoprotein function: Recent advances. J Am Coll Cardiol 2005;46:1792–8. https://doi.org/10.1016/j.jacc.2005.06.080.

[3] Khovidhunkit W, Duchateau PN, Medzihradszky KF, Moser AH, Naya-Vigne J, Shigenaga JK, et al. Apolipoproteins A-IV and A-V are acute-phase proteins in mouse HDL. Atherosclerosis 2004;176:37–44. https://doi.org/10.1016/j.atherosclerosis.2004.04.018.

[4] Laskowitz DT, Lee DM, Schmechel D, Staats HF. Altered immune responses in apolipoprotein E-deficient mice. J Lipid Res 2000;41:613–20. https://doi.org/10.1016/s0022-2275(20)32409-3.

[5] Vaisar T, Pennathur S, Green PS, Gharib SA, Hoofnagle AN, Cheung MC, et al. Shotgun proteomics implicates protease inhibition and complement activation in the antiinflammatory properties of HDL. J Clin Invest 2007;117:746–56. https://doi.org/10.1172/JCI26206.

[6] Kailemia MJ, Wei W, Nguyen K, Beals E, Sawrey-Kubicek L, Rhodes C, et al. Targeted Measurements of O- and N-Glycopeptides Show That Proteins in High Density Lipoprotein Particles Are Enriched with Specific Glycosylation Compared to Plasma. J Proteome Res 2018;17:834–45. https://doi.org/10.1021/acs.jproteome.7b00604.

[7] Huang J, Lee H, Zivkovic AM, Smilowitz JT, Rivera N, German JB, et al. Glycomic analysis of high density lipoprotein shows a highly sialylated particle. J Proteome Res 2014;13:681–91. https://doi.org/10.1021/pr4012393.

[8] Schindler PA, Settineri CA, Collet X, Fielding CJ, Burlingame AL. Site-specific detection and structural characterization of the glycosylation of human plasma proteins lecithin:cholesterol acyltransferase and apolipoprotein D using HPLC/electrospray mass spectrometry and sequential glycosidase digestion. Protein Sci 1995;4:791–803. https://doi.org/10.1002/pro.5560040419.

[9] Zhu C, Wong M, Li Q, Sawrey-Kubicek L, Beals E, Rhodes CH, et al. Site-Specific Glycoprofiles of HDL-Associated ApoE are Correlated with HDL Functional Capacity and Unaffected by Short-Term Diet. J Proteome Res 2019;18:3977–84. https://doi.org/10.1021/acs.jproteome.9b00450.

[10] Remaley AT, Wong AW, Schumacher UK, Meng MS, Brewer HB, Hoeg JM. O-linked glycosylation modifies the association of apolipoprotein A-II to high density lipoproteins. J Biol Chem 1993;268:6785–90. https://doi.org/10.1016/s0021-9258(18)53318-4.

[11] Vaisar T. Proteomics Investigations of HDL: Challenges and Promise. Curr Vasc Pharmacol 2012;10:410–21. https://doi.org/10.2174/157016112800812755.

[12] Li Q, Kailemia MJ, Merleev AA, Xu G, Serie D, Danan LM, et al. Site-Specific Glycosylation Quantitation of 50 Serum Glycoproteins Enhanced by Predictive Glycopeptidomics for Improved Disease Biomarker Discovery. Anal Chem 2019;91:5433–45. https://doi.org/10.1021/acs.analchem.9b00776.

[13] Hong Q, Ruhaak LR, Stroble C, Parker E, Huang J, Maverakis E, et al. A Method for Comprehensive Glycosite-Mapping and Direct Quantitation of Serum Glycoproteins. J Proteome Res 2015;14:5179–92. https://doi.org/10.1021/acs.jproteome.5b00756.

[14] Zheng JJ, Agus JK, Hong B V., Tang X, Rhodes CH, Houts HE, et al. Isolation of HDL by sequential flotation ultracentrifugation followed by size exclusion chromatography reveals size-based enrichment of HDL-associated proteins. Sci Rep 2021;11. https://doi.org/10.1038/s41598-021-95451-3.

[15] Adu-Afarwuah S, Lartey A, Okronipa H, Ashorn P, Peerson JM, Arimond M, et al. Small-quantity, lipid-based nutrient supplements provided to women during pregnancy and 6 mo postpartum and to their infants from 6 mo of age increase the mean attained length of 18-mo-old children in semi-urban Ghana: A randomized controlled trial. Am J Clin Nutr 2016;104:797–808. https://doi.org/10.3945/ajcn.116.134692.

[16] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47. https://doi.org/10.1093/nar/gkv007.

[17] Hong B V., Zhu C, Wong M, Sacchi R, Rhodes CH, Kang JW, et al. Lipid-Based Nutrient Supplementation Increases High-Density Lipoprotein (HDL) Cholesterol Efflux Capacity and Is Associated with Changes in the HDL Glycoproteome in Children. ACS Omega 2021;6:32022–31. https://doi.org/10.1021/acsomega.1c04811.

[18] Abdelhamid HN, Wu HF. Proteomics analysis of the mode of antibacterial action of nanoparticles and their interactions with proteins. TrAC - Trends Anal Chem 2015;65:30–46. https://doi.org/10.1016/j.trac.2014.09.010.

[19] Wu HF, Gopal J, Abdelhamid HN, Hasan N. Quantum dot applications endowing novelty to analytical proteomics. Proteomics 2012;12:2949–61. https://doi.org/10.1002/pmic.201200295.

[20] Chen ZY, Abdelhamid HN, Wu HF. Effect of surface capping of quantum dots (CdTe) on proteomics. Rapid Commun Mass Spectrom 2016;30:1403–12. https://doi.org/10.1002/rcm.7575.

[21] Jong MC, Hofker MH, Havekes LM. Role of apoCs in lipoprotein metabolism: Functional differences between ApoC1, ApoC2, and ApoC3. Arterioscler Thromb Vasc Biol 1999;19:472–84. https://doi.org/10.1161/01.ATV.19.3.472.

[22] Yamazaki Y, Zhao N, Caulfield TR, Liu CC, Bu G. Apolipoprotein E and Alzheimer disease: pathobiology and targeting strategies. Nat Rev Neurol 2019;15:501–18. https://doi.org/10.1038/s41582-019-0228-7.

[23] Marais AD. Apolipoprotein E in lipoprotein metabolism, health and cardiovascular disease. Pathology 2019;51:165–76. https://doi.org/10.1016/j.pathol.2018.11.002.

[24] Li X, Ma Y, Wei X, Li Y, Wu H, Zhuang J, et al. Clusterin in Alzheimer's disease: A player in the biological behavior of amyloid-beta. Neurosci Bull 2014;30:162–8. https://doi.org/10.1007/s12264-013-1391-2.

[25] Merlotti A, Malizia AL, Michea P, Bonte PE, Goudot C, Carregal MS, et al. Aberrant fucosylation enables breast cancer clusterin to interact with dendritic cell-specific ICAM-grabbing non-integrin (DC-SIGN). Oncoimmunology 2019;8. https://doi.org/10.1080/2162402X.2019.1629257.

[26] Kluve-Beckerman B, Long GL, Benson MD. DNA sequence evidence for polymorphic forms of human serum amyloid A (SAA). Biochem Genet 1986;24:795–803. https://doi.org/10.1007/BF00554519.

[27] de Serres F, Blanco I. Role of alpha-1 antitrypsin in human health and disease. J Intern Med 2014;276:311–35. https://doi.org/10.1111/joim.12239.