



## OPEN Transcriptomic and glycomic analyses highlight pathway-specific glycosylation alterations unique to Alzheimer's disease

Xinyu Tang<sup>1,6</sup>, Jennyfer Tena<sup>2,6</sup>, Jacopo Di Lucente<sup>3,4</sup>, Izumi Maezawa<sup>3,4</sup>, Danielle J. Harvey<sup>5</sup>, Lee-Way Jin<sup>3,4</sup>, Carlito B. Lebrilla<sup>2</sup> & Angela M. Zivkovic<sup>1</sup>✉

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. 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). 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 has-miR-1-3p and has-miR-16-5p, respectively. 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.

### Abbreviations

AD	Alzheimer's disease
MCI	Mild cognitive impairment
NCI	No cognitive impairment
MGAT1	Alpha-1,3-mannosyl-glycoprotein 2-beta- <i>N</i> -acetylglucosaminyltransferase
B4GALT	β-1,4-Galactosyltransferase 1
GALNT	Polypeptide <i>N</i> -acetylgalactosaminyltransferase
ST6GALNAC	Alpha- <i>N</i> -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

<sup>1</sup>Department of Nutrition, University of California, Davis, One Shields Ave, Davis, CA 95616, USA. <sup>2</sup>Department of Chemistry, University of California, Davis, Davis, CA, USA. <sup>3</sup>Department of Pathology and Laboratory Medicine, School of Medicine, University of California, Davis, Sacramento, CA, USA. <sup>4</sup>UC Davis MIND Institute, Sacramento, CA, USA. <sup>5</sup>Division of Biostatistics, Department of Public Health Sciences, School of Medicine, University of California, Davis, Davis, CA, USA. <sup>6</sup>These authors contributed equally: Xinyu Tang and Jennyfer Tena. ✉email: amzivkovic@ucdavis.edu

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
A $\beta$	Amyloid $\beta$
APP	Amyloid precursor protein
MSBB	Mount Sinai Brain Bank
ROSMAP	Religious orders study/memory and aging project

Alzheimer's disease (AD) is the most common form of dementia in the elderly, characterized by the accumulation of extracellular amyloid- $\beta$  (A $\beta$ ) and intracellular hyperphosphorylated tau, as well as the accumulation of cholesterol-containing lipid droplets<sup>1,2</sup>. However, treatments to reduce A $\beta$  load have failed to stop AD progression<sup>3,4</sup>, 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<sup>5-9</sup>. 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<sup>10</sup>. Similarly, dysregulated N-glycosylation and sialylation patterns in astrocytes and microglia affect extracellular functions and inflammatory responses<sup>10,11</sup>.

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<sup>12</sup>. 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<sup>13</sup>. 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 A $\beta$  production by reducing O-GalNAc formation of the amyloid precursor protein (APP)<sup>14</sup>. In contrast, GALNT6 enzymatic activity caused reduced A $\beta$ 40 and A $\beta$ 42 generation with excess GalNAc-type O-glycosylation of APP<sup>15</sup>. Although several studies have explored the brain glycoproteome in AD<sup>10,16</sup>, 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 Supplementary Tables S1, S2 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<sup>17</sup> 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 infor-

mation 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 and Hochberg (BH) with a threshold of adjusted p-value  $\leq 0.05$  (Supplementary Methods).

**Glycosylation-related gene selection.** A list of 219 glycosylation-related genes was obtained from a previous publication<sup>18</sup>.

**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<sup>19</sup>. 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  $-80\text{ }^{\circ}\text{C}$  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*-GalNAc-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 Supplementary Table S3. Gene expression was normalized to an endogenous reference gene,  $\beta$ -actin. Data were analyzed by the  $2^{-\Delta\Delta C_t}$  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) (Supplementary Table S4) were analyzed. Tissue samples were homogenized, resuspended and lysed. Lysates were pelleted by centrifugation at  $2000\times g$  for 10 min to remove the nuclear fraction and cell debris. The supernatant was centrifuged at  $200,000\times g$ ,  $4\text{ }^{\circ}\text{C}$  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  $2\text{ }\mu\text{L}$  ( $500,000$  units/mL) of peptide *N*-glycosidase F (glycerol-free; New England BioLabs, cat. no. P0705L). Ultracentrifugation at  $200,000\times g$  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  $30\text{ }\mu\text{L}$  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<sup>20</sup>. We used the motif collection based on human hg38 genome assembly and RefSeq genes and searched 10 kbp 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<sup>21</sup>. (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<sup>22</sup> as shown here: glycan abundance  $\sim$  diagnosis  $\times$  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.

**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.

## 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 (Supplementary Table S5). The number of DEGs in each brain region varied with approximately 13,822 protein-coding genes differentially expressed in at least one brain region (adjusted  $p$ -value  $< 0.05$ , Supplementary Fig. 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 (Supplementary Fig. S1B). Two upregulated genes, including alpha-*N*-acetylgalactosaminide alpha-2,6-sialyltransferase 2 (ST6GALNAC2), and three downregulated genes, including ST6GALNAC5 were observed in all brain regions (Supplementary Fig. 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 (Supplementary Fig. 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 (Supplementary Fig. 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 (Fig. 1A, Supplementary Fig. 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 fourfold increase in gene expression in AD participants vs. healthy controls (Fig. 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 (Fig. 1C). The percent abundances of two intermediates of the pathway, the direct substrate for the MGAT1 enzyme, Hex:5 HexNAc:2 (Intermediate 1 in Fig. 1D,E), and the direct product of the MGAT1 enzyme, Hex:5 HexNAc:3 (Intermediate 2 in Fig. 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 (Supplementary Fig. S2B–D).

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

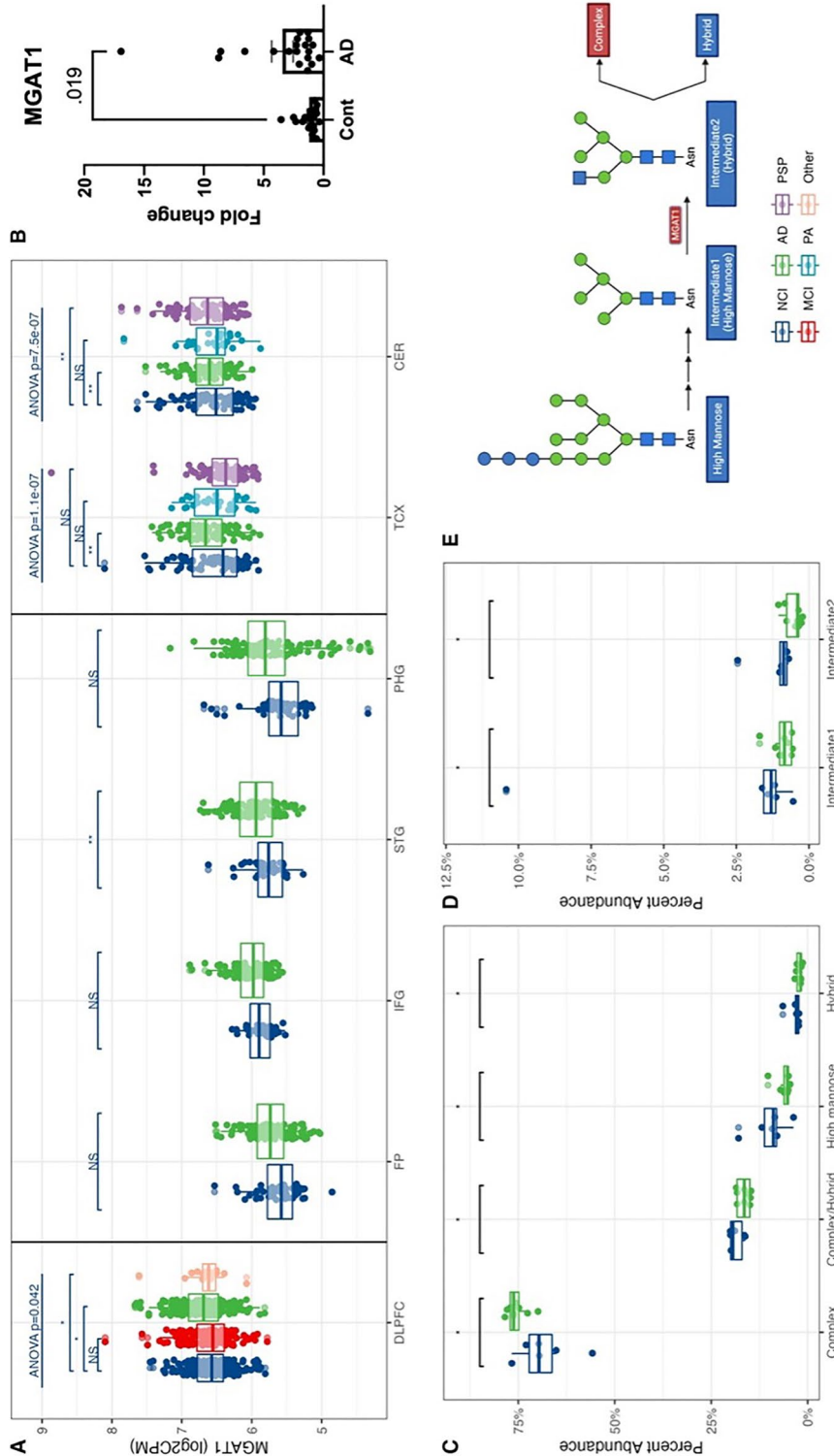
The glycosylation process also involves elongation and branching biosynthetic steps. However, most glycosyltransferases for these steps belong to isoenzyme families with poorly understood non-redundant functions. We detected 17 glycosyltransferases for elongation and branching (Supplementary Fig. S3A).  $\beta$ -1,4-galactosyltransferase 1 (B4GALT1) adds terminal galactose to N-glycans to generate type 2 LacNAc units (Fig. 2D). The gene expression of B4GALT1 was substantially increased in AD participant brains in TCX (Fig. 2A, Supplementary Fig. S3A) and was not significantly changed in PA and PSP individuals, indicating that B4GALT1 may be an AD-specific target gene (Fig. 2A). The increased expression of B4GALT1 was also observed by qPCR in an independent set of tissue samples (Fig. 2B), showing approximately a twofold 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 (Fig. 2C). Meanwhile, the proportion of N-glycans with six galactoses significantly increased in the MTC (Supplementary Fig. 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 (Supplementary Fig. 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 (Supplementary Fig. S3B,C).  $\beta$ -1,3-GalTs generate type-1 LacNAc units, which are relatively high in O-glycans and glycosphingolipids of humans.

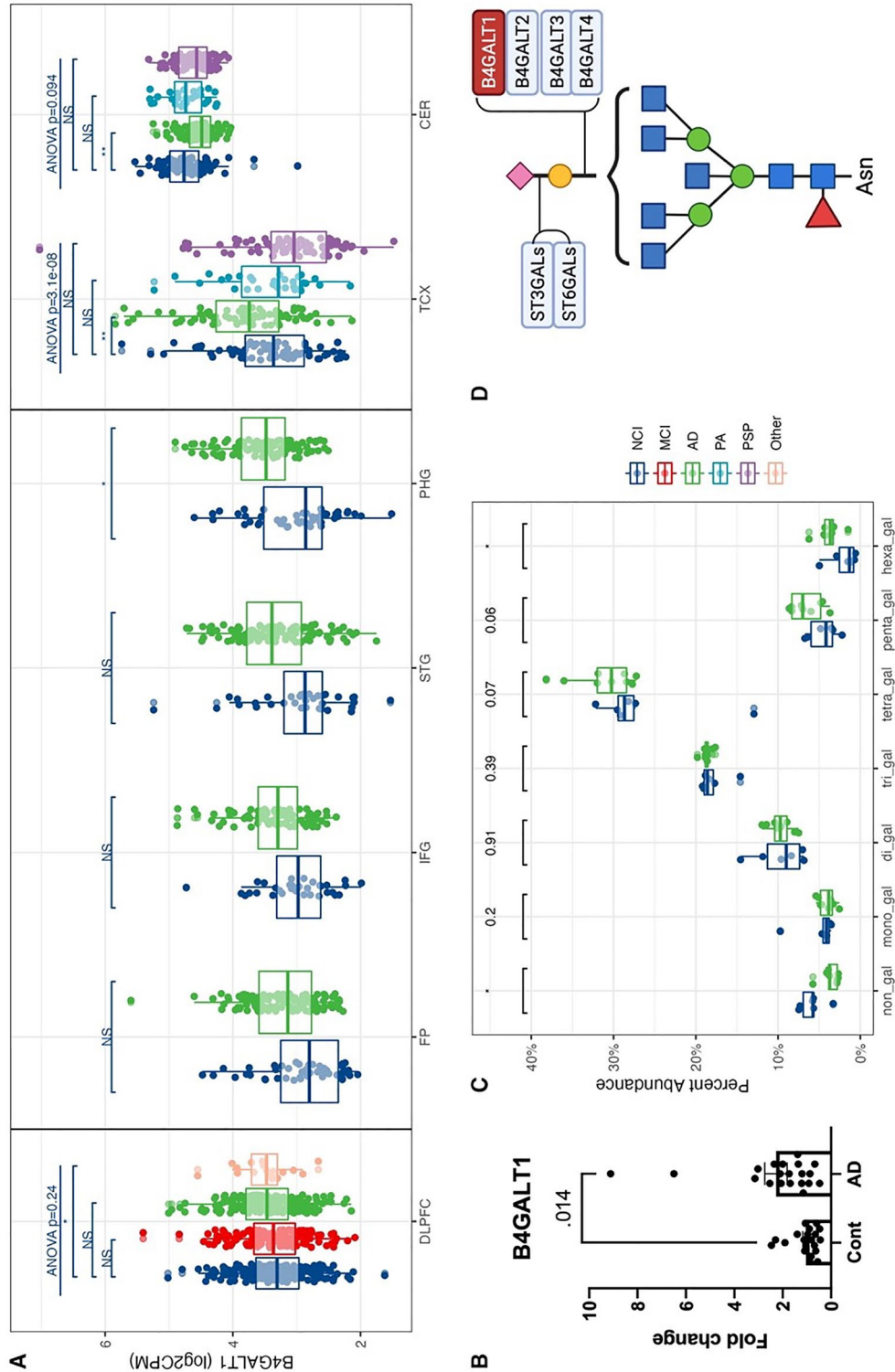
**Heterogeneous differential expression of isoenzymes involved in O-GalNAc-glycan initiation and alpha-2,6-sialylation on GalNAc residues.**

We observed heterogeneous differential expression within several notable isoenzyme families. For example, in the GALNT family of enzymes that initiate O-GalNAc-glycan biosynthesis (Fig. 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 (Fig. 3A,C, Supplementary Fig. 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 (Fig. 3A,C, Supplementary Fig. S4A–C). qPCR-based expression

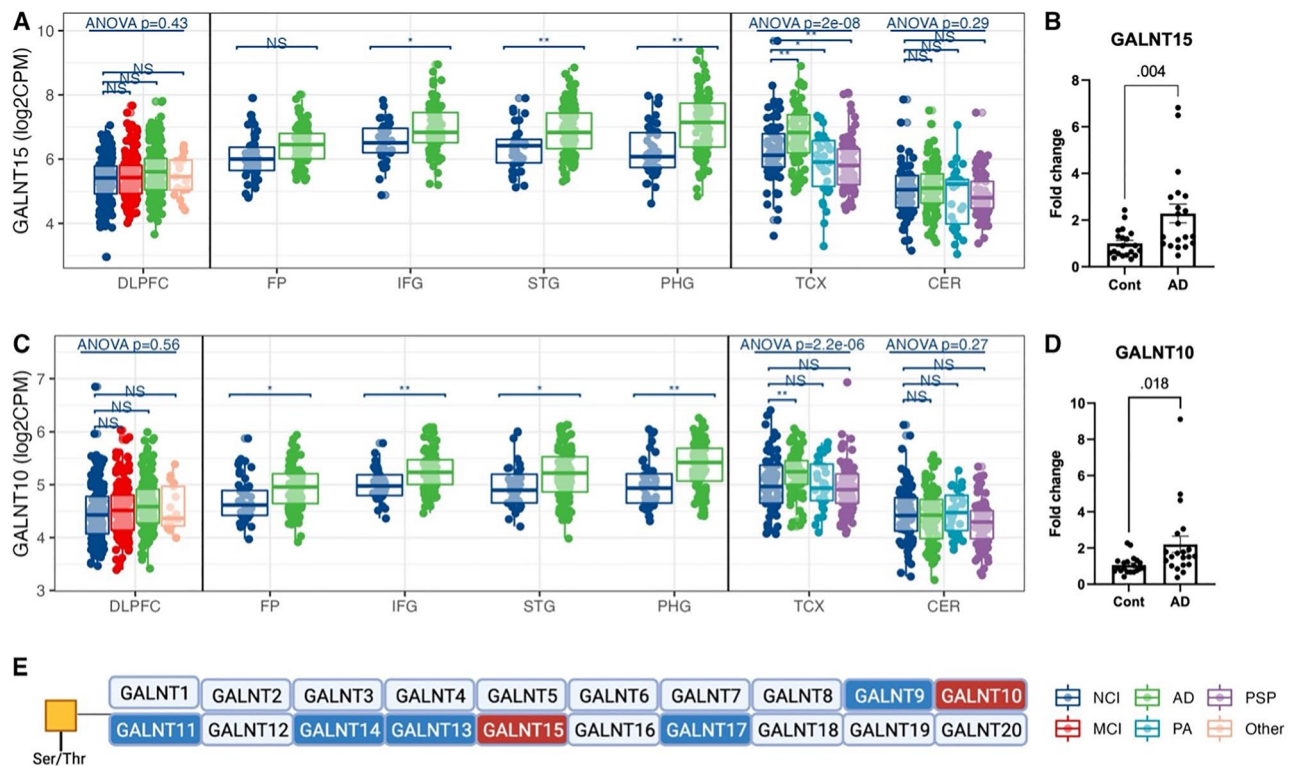




**Figure 1.** Increased MGAT1 expression corroborated by a raised complex and reduced high-mannose N-glycans abundance in AD. (A) MGAT1 gene expression from RNA-seq across human brains with different diagnoses. Data presented as log<sub>2</sub> count per million (log<sub>2</sub>CPM). Data from different 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, \*\*p-value < 0.01, \*\*\*p-value < 0.001. ANOVA p-values were after the multiple testing correction over all protein-coding genes. *DLPFC* dorsolateral prefrontal cortex, *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 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. DLFPFC 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 galactosylated and sialylated 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 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 log<sub>2</sub> count per million (log<sub>2</sub>CPM). 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).

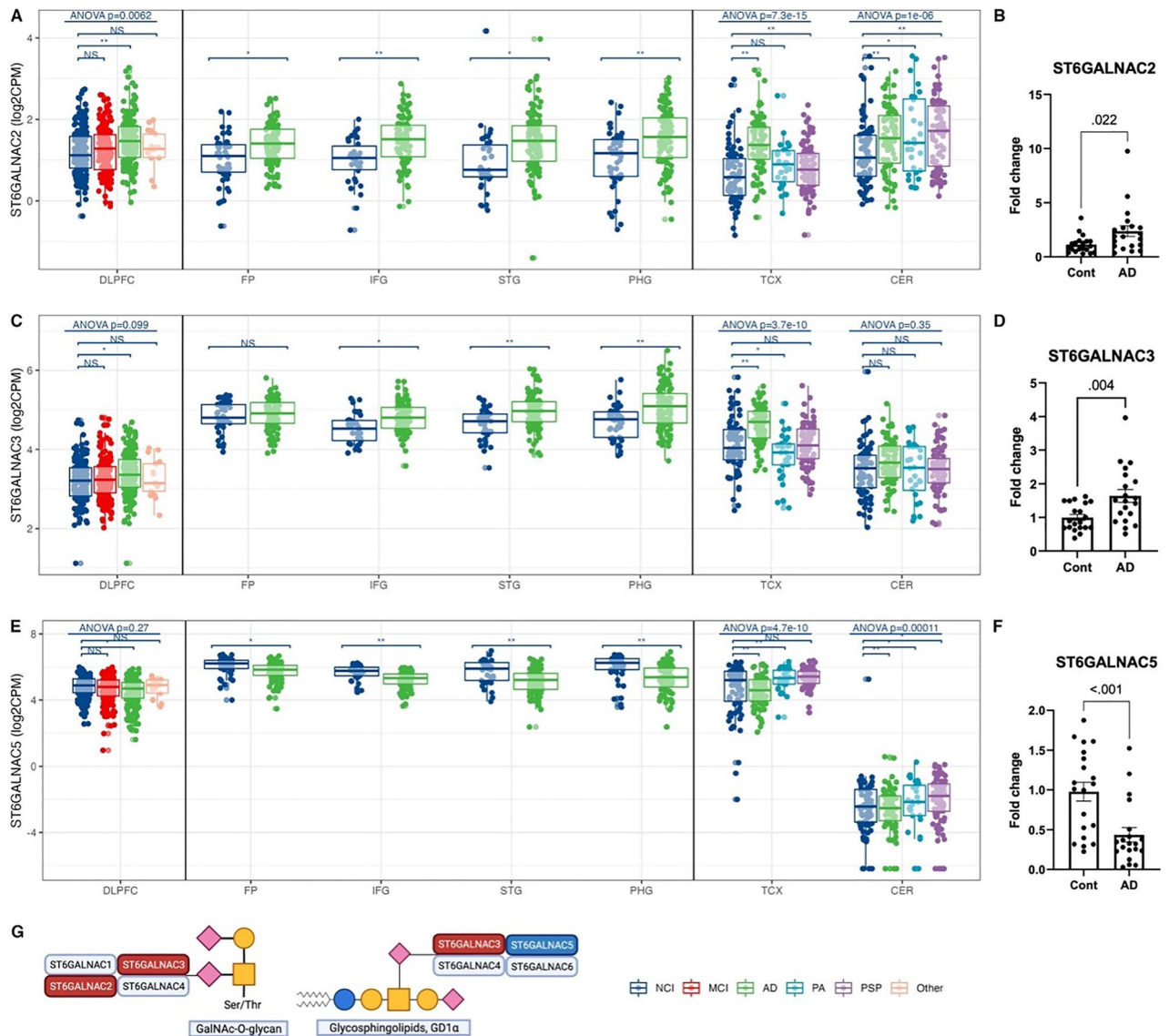
analysis in an independent sample set of tissues showed a twofold increase in gene expression of GALNT15 and GALNT10 but not GALNT11 in AD participants compared with controls (Fig. 3B,D, Supplementary Fig. S4D).

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 (Fig. 4G), ST6GALNAC2 and ST6GALNAC3 expression increased in AD participants, while ST6GALNAC5 decreased (Fig. 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 (Fig. 4A–F). Notably, the expression level of ST6GALNAC5 was much lower in the CER than in the six cortical regions across all diagnosis groups (Fig. 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 (Supplementary Fig. 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. According to the RNA-seq results, the temporal regions had more altered genes than the frontal regions (Figure 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 (Fig. 5A,D, Supplementary Fig. S6A). qPCR results confirmed a twofold increase in the expression of these genes in AD participants (Fig. 5B,E). Other critical genes were also upregulated in AD, such as PIGV (Supplementary Fig. 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



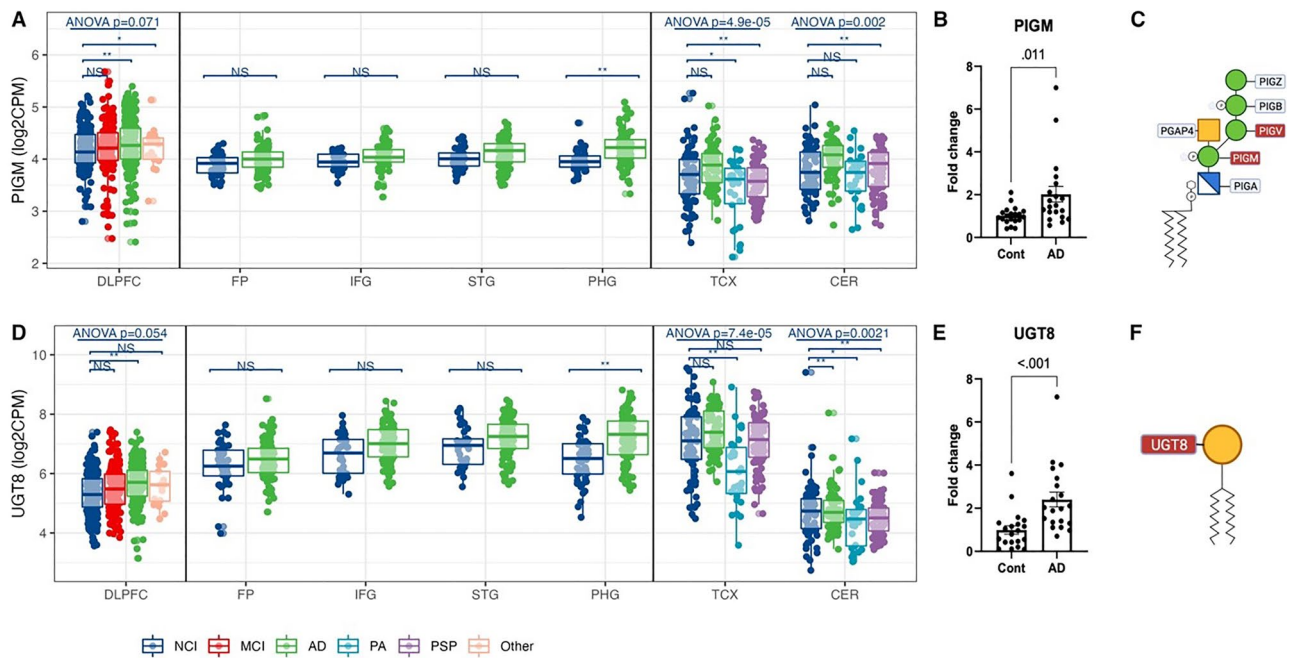


**Figure 4.** Differential expression of alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase (ST6GALNAC) gene family in AD across brain regions. **(A)** ST6GALNAC2 gene expression from RNA-seq across human brains with different diagnoses. Data presented as log2 count per million (log2CPM). Data 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).

in the brains of AD participants (Fig. 5C). Of the glycosphingolipid biosynthesis pathways, since UGT8 catalyzes a crucial step of galactocerebroside biosynthesis by transferring galactose to ceramide (Fig. 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) (Supplementary Fig. S6B,D).





**Figure 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 log<sub>2</sub> count per million (log<sub>2</sub>CPM). 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).

**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 (Supplementary Table S6, Supplementary Fig. 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 (Supplementary Table S6, Supplementary Fig. S7B).

Using databases containing experimentally validated miRNA-target interactions, we identified miRNAs that possibly regulate various glycosylation pathways (Supplementary Table 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 (Supplementary Table S7, Supplementary Fig. 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 (Supplementary Table S7, Supplementary Fig. 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 (Supplementary Fig. 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<sup>23,24</sup>. 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 (Supplementary Fig. S9). A previous analysis using transcriptomic data from the ROSMAP, MSBB, and Mayo data sets revealed 4485 DEGs with a fixed meta-analysis model<sup>25</sup>.

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<sup>15</sup>, while another study found upregulation of STT3B, MAGT1, ALG2/6/12, and downregulation of RPN2, DAD1, OST4, OSTC, TUSC3, ALG3/9/12 in AD<sup>10</sup>.

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<sup>26</sup>. 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<sup>13,27</sup>. 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<sup>13</sup>. 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<sup>28</sup>. However, other studies using different glycomic analytical approaches have reported the most abundant structures found were high mannose structures<sup>7,29</sup>.

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<sup>30</sup>. Conversely, B4GALT1 knockdown in mice resulted in an inhibition of glioblastoma development and increased survival<sup>31</sup>. 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<sup>32,33</sup>. 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<sup>34,35</sup>. 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<sup>29</sup>. 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 et al. conducted a glycoproteomic study in which they compared their *N*-glycoproteomic data to four different transcriptome datasets for brain tissues<sup>10</sup>. 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<sup>36–38</sup>, 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<sup>39–43</sup>, 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 sialyltransferases 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<sup>44</sup>. 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<sup>45,46</sup>. CD33 overexpression in the human brain is associated with impaired

cognition<sup>45,47,48</sup>. When alpha-2,3 and, preferentially, alpha-2,6 sialylated glycans bind CD33, the A $\beta$  phagocytosis activity of microglia increases<sup>11,49,50</sup>. 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 DLPCF 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<sup>51,52</sup>. 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<sup>52,53</sup>. The changes in glycosphingolipid patterns are cell-specific as glucosylceramide-derived gangliosides are enriched in neurons, and galactosylceramides are enriched in oligodendrocytes and myelin<sup>52–54</sup>. 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<sup>55,56</sup>. 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<sup>57,58</sup>. Its expression impaired spatial learning and memory in rats and could be induced by A $\beta$  in a dose-dependent manner<sup>59</sup>. 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<sup>60</sup>. Moreover, the expression of hsa-miR-124 gradually decreased in AD, which led to up-regulated BACE1 expression, ultimately responsible for A $\beta$  production<sup>61,62</sup>. hsa-miR-34a has been reported to regulate ALG13, FUT8, GALNT7, and ST3GAL5 in neuroblastoma and hepatocellular carcinoma<sup>63–66</sup>. 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<sup>67</sup>. In addition, hsa-miR-34a-5p has been found to be dysregulated in AD and other neurodegenerative diseases<sup>68</sup>. hsa-miR-16-5p was upregulated by A $\beta$  in AD and induced neuronal apoptosis<sup>69</sup>. 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<sup>70</sup>. 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<sup>71</sup>. 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. glutamergic 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<sup>72</sup>. 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.

## Data availability

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-glycomics datasets generated and analyzed are included in the supplementary information files of this article. Mass spectrometry data for glycomics (<https://doi.org/10.25345/C5J67920T>, MSV000090677) is available on MassIVE data repository upon reasonable request (contact: xctang@ucdavis.edu).

Received: 5 January 2023; Accepted: 8 May 2023

Published online: 15 May 2023

## References

- Jack, C. R. *et al.* NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's Dement.* **14**, 535–562 (2018).
- Stelzmann, R. A., Norman Schnitzlein, H. & Reed, M. F. An english translation of alzheimer's 1907 paper, "über eine eigenartige erkrankung der hirnrinde". *Clin. Anat.* **8**, 429–431 (1995).
- Pahnke, J., Walker, L. C., Scheffler, K. & Krohn, M. Alzheimer's disease and blood–brain barrier function—Why have anti- $\beta$ -amyloid therapies failed to prevent dementia progression?. *Neurosci. Biobehav. Rev.* **33**, 1099–1108 (2009).
- Soejitno, A., Tjan, A. & Purwata, T. E. Alzheimer's disease: Lessons learned from amyloidocentric clinical trials. *CNS Drugs* **29**, 487–502 (2015).
- Klarić, T. S. & Lauc, G. The dynamic brain N-glycome. *Glycoconj. J.* <https://doi.org/10.1007/s10719-022-10055-x> (2022).
- Frenkel-Pinter, M. *et al.* Interplay between protein glycosylation pathways in Alzheimer's disease. *Sci. Adv.* **3**, e1601576 (2017).
- Gizaw, S. T., Ohashi, T., Tanaka, M., Hinou, H. & Nishimura, S. I. Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. *Biochim. Biophys. Acta Gen. Subj.* **1860**, 1716–1727 (2016).
- Sihlbom, C., Davidsson, P., Sjögren, M., Wahlund, L. O. & Nilsson, C. L. Structural and quantitative comparison of cerebrospinal fluid glycoproteins in Alzheimer's disease patients and healthy individuals. *Neurochem. Res.* **33**, 1332–1340 (2008).
- Kanninen, K., Goldsteins, G., Auriola, S., Alafuzoff, I. & Koistinaho, J. Glycosylation changes in Alzheimer's disease as revealed by a proteomic approach. *Neurosci. Lett.* **367**, 235–240 (2004).
- Zhang, Q., Ma, C., Chin, L. S. & Li, L. Integrative glycoproteomics reveals protein n-glycosylation aberrations and glycoproteomic network alterations in Alzheimer's disease. *Sci. Adv.* **6**, eabc5802 (2020).
- Estus, S. *et al.* Evaluation of CD33 as a genetic risk factor for Alzheimer's disease. *Acta Neuropathol.* **138**, 187–199 (2019).
- Freeze, H. H., Eklund, E. A., Ng, B. G. & Patterson, M. C. Neurological aspects of human glycosylation disorders. *Annu. Rev. Neurosci.* **38**, 105–125 (2015).
- Ye, Z. & Marth, J. D. N-glycan branching requirement in neuronal and postnatal viability. *Glycobiology* **14**, 547–558 (2004).
- Liu, F. *et al.* The small molecule luteolin inhibits N-acetyl- $\alpha$ -galactosaminyltransferases and reduces mucin-type O-glycosylation of amyloid precursor protein. *J. Biol. Chem.* **292**, 21304–21319 (2017).
- Akasaka-Manya, K. *et al.* Excess APP O-glycosylation by GalNac-T6 decreases Ab production. *J. Biochem.* **161**, 99–111 (2017).
- Fang, P. *et al.* Multilayered N-glycoproteome profiling reveals highly heterogeneous and dysregulated protein N-glycosylation related to Alzheimer's disease. *Anal. Chem.* **92**, 867–874 (2020).
- McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
- Schjoldager, K. T., Narimatsu, Y., Joshi, H. J. & Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 729–749 (2020).
- Montine, T. J. *et al.* National institute on aging—Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease: A practical approach. *Acta Neuropathol.* **123**, 1–11 (2012).
- Aibar, S. *et al.* SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods.* **14**, 1083–1086 (2017).
- Ru, Y. *et al.* The multiMiR R package and database: Integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res.* **42**, e133 (2014).
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).
- Guennewig, B. *et al.* Defining early changes in Alzheimer's disease from RNA sequencing of brain regions differentially affected by pathology. *Sci. Rep.* **11**, 1–15 (2021).
- Williams, J. B., 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.* **3**, fcab123 (2021).
- Wan, Y. W. *et al.* Meta-analysis of the Alzheimer's disease human brain transcriptome and functional dissection in mouse models. *Cell Rep.* **32**, 107908 (2020).
- Cho, B. G., Veillon, L. & Mechref, Y. N-glycan profile of cerebrospinal fluids from Alzheimer's disease patients using liquid chromatography with mass spectrometry. *J. Proteome Res.* **18**, 3770–3779 (2019).
- 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.* **91**, 728–732 (1994).
- Gaunitz, S., Tjernberg, L. O. & Schedin-Weiss, S. The N-glycan profile in cortex and hippocampus is altered in Alzheimer disease. *J. Neurochem.* **159**, 292–304 (2021).
- Williams, S. E. *et al.* Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. *Nat. Commun.* **13**, 275 (2022).



30. Niu, S. *et al.* Altered  $\beta$ -1,4-galactosyltransferase I expression during early inflammation after spinal cord contusion injury. *J. Chem. Neuroanat.* **35**, 245–256 (2008).
31. Wang, P., Li, X. & Xie, Y. B4GalT1 regulates apoptosis and autophagy of glioblastoma in vitro and in vivo. *Technol. Cancer Res. Treat.* **19**, 153303382098010 (2020).
32. Narimatsu, Y. *et al.* An atlas of human glycosylation pathways enables display of the human glycome by gene engineered cells. *Mol. Cell.* **75**, 394–407.e5 (2019).
33. Stanley, P. What have we learned from glycosyltransferase knockouts in mice?. *J. Mol. Biol.* **428**, 3166–3182 (2016).
34. Huang, Y. F. *et al.* Global mapping of glycosylation pathways in human-derived cells. *Dev. Cell.* **56**, 1195–1209.e7 (2021).
35. Dworkin, L. A., Clausen, H. & Joshi, H. J. Applying transcriptomics to study glycosylation at the cell type level. *iScience.* **25**, 104419 (2022).
36. Steentoft, C. *et al.* Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. *Nat. Methods.* **8**, 977–982 (2011).
37. de Las, R. M., Lira-Navarrete, E., Gerken, T. A. & Hurtado-Guerrero, R. Polypeptide GalNAc-Ts: From redundancy to specificity. *Curr. Opin. Struct. Biol.* **56**, 87–96 (2019).
38. Schjoldager, K. T. B. G. *et al.* Probing isoform-specific functions of polypeptide GalNAc-transferases using zinc finger nuclease glycoengineered SimpleCells. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 9893–9898 (2012).
39. Cheng, L. *et al.* Characterization of a novel human UDP-GalNAc transferase, pp-GalNAc-T10. *FEBS Lett.* **531**, 115–121 (2002).
40. Raman, J., Guan, Y., Perrine, C. L., Gerken, T. A. & Tabak, L. A. UDP-N-acetyl  $\alpha$ -D-galactosamine: Polypeptide N-acetylgalactosaminyltransferases: Completion of the family tree. *Glycobiology* **22**, 768–777 (2012).
41. Lee, Y. C. *et al.* Molecular cloning and functional expression of two members of mouse NeuAc $\alpha$ 2,3Gal $\beta$ 1,3GalNAc GalNAc $\alpha$ 2,6-sialyltransferase family, ST6GalNAc III and IV. *J. Biol. Chem.* **274**, 11958–11967 (1999).
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.* **269**, 19048–19053 (1994).
43. Okajima, T. *et al.* Molecular cloning of brain-specific GD1a synthase (ST6GalNAc V) containing CAG/glutamine repeats. *J. Biol. Chem.* **274**, 30557–30562 (1999).
44. Wißfeld, J. *et al.* Deletion of Alzheimer's disease-associated CD33 results in an inflammatory human microglia phenotype. *Glia* **69**, 1393–1412 (2021).
45. Naj, A. C. *et al.* Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat. Genet.* **43**, 436–443 (2011).
46. Malik, M. *et al.* CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. *J. Neurosci.* **33**, 13320–13325 (2013).
47. Gong, L. *et al.* The CD33 genotype associated cognitive performance was bidirectionally modulated by intrinsic functional connectivity in the Alzheimer's disease spectrum. *Biomed. Pharmacother.* **115**, 108903 (2019).
48. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
49. Varki, A. & Angata, T. Siglecs—The major subfamily of I-type lectins. *Glycobiology* **16**, 1–27 (2006).
50. Bhattacharjee, A. *et al.* Increasing phagocytosis of microglia by targeting CD33 with liposomes displaying glycan ligands. *J. Control Release.* **338**, 680–693 (2021).
51. Rocha, E. M. *et al.* Progressive decline of glucocerebrosidase in aging and Parkinson's disease. *Ann. Clin. Transl. Neurol.* **2**, 433–438 (2015).
52. Belarbi, K. *et al.* Glycosphingolipids and neuroinflammation in Parkinson's disease. *Mol. Neurodegener.* **15**, 59 (2020).
53. Sandhoff, R. & Sandhoff, K. Emerging concepts of ganglioside metabolism. *FEBS Lett.* **592**, 3835–3864 (2018).
54. Norton, W. T. & Autilio, L. A. The lipid composition of purified bovine brain myelin. *J. Neurochem.* **13**, 213–222 (1966).
55. Tan, R. P. A., Leshchyn'Ksa, I. & Sytnyk, V. Glycosylphosphatidylinositol-anchored immunoglobulin superfamily cell adhesion molecules and their role in neuronal development and synapse regulation. *Front. Mol. Neurosci.* <https://doi.org/10.3389/fnmol.2017.00378> (2017).
56. Puzzo, D. *et al.* F3/Contactin promotes hippocampal neurogenesis, synaptic plasticity, and memory in adult mice. *Hippocampus* **23**, 1367–1382 (2013).
57. Rangaraju, S. *et al.* Identification and therapeutic modulation of a pro-inflammatory subset of disease-associated-microglia in Alzheimer's disease. *Mol. Neurodegener.* **13**, 1–25 (2018).
58. Butturini, E., Boriero, D., Carcereri de Prati, A. & Mariotto, S. STAT1 drives M1 microglia activation and neuroinflammation under hypoxia. *Arch. Biochem. Biophys.* **669**, 22–30 (2019).
59. Hsu, W. L., Ma, Y. L., Hsieh, D. Y., Liu, Y. C. & Lee, E. H. STAT1 negatively regulates spatial memory formation and mediates the memory-impairing effect of A $\beta$ . *Neuropsychopharmacology* **39**, 746–758 (2014).
60. Liu, Y. X. *et al.* MIR-124-3p/B4GALT1 axis plays an important role in SOCS3-regulated growth and chemo-sensitivity of CML. *J. Hematol. Oncol.* **9**, 1–2 (2016).
61. Fang, M. *et al.* The miR-124 regulates the expression of BACE1/ $\beta$ -secretase correlated with cell death in Alzheimer's disease. *Toxicol. Lett.* **209**, 94–105 (2012).
62. An, F. *et al.* MiR-124 acts as a target for Alzheimer's disease by regulating BACE1. *Oncotarget* **8**, 114065–114071 (2017).
63. Cai, H. *et al.* MiRNA expression profiles reveal the involvement of miR-26a, miR-548l and miR-34a in hepatocellular carcinoma progression through regulation of ST3GAL5. *Lab Invest.* **97**, 530–542 (2017).
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.* **9**, 1293–1298 (2014).
65. Bernardi, C., Soffientini, U., Piacente, F. & Tonetti, M. G. Effects of microRNAs on fucosyltransferase 8 (FUT8) expression in hepatocarcinoma cells. *PLoS ONE* **8**, e76540 (2013).
66. De Antonellis, P. *et al.* Early targets of miR-34a in neuroblastoma. *Mol. Cell Proteomics.* **13**, 2114–2131 (2014).
67. Chua, C. E. L. & Tang, B. L. miR-34a in neurophysiology and neuropathology. *J. Mol. Neurosci.* <https://doi.org/10.1007/s12031-018-1231-y> (2019).
68. van den Berg, M. M. J. *et al.* Circulating microRNAs as potential biomarkers for psychiatric and neurodegenerative disorders. *Prog. Neurobiol.* **185**, 101732 (2020).
69. Kim, Y. J. *et al.* miR-16-5p is upregulated by amyloid  $\beta$  deposition in Alzheimer's disease models and induces neuronal cell apoptosis through direct targeting and suppression of BCL-2. *Exp. Gerontol.* **136**, 110954 (2020).
70. Everaert, C. *et al.* Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Sci. Rep.* **7**, 1559 (2017).
71. Song, L. *et al.* STAB: A spatio-temporal cell atlas of the human brain. *Nucleic Acids Res.* **49**, D1029–D1037 (2021).
72. Williams, S. E. *et al.* Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. *Nat. Commun.* <https://doi.org/10.1038/s41467-021-27781-9> (2022).

## 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, R01AG15819, R01AG17917, R01AG036836, U01AG46152, U01AG61356, U01AG046139, P50AG016574, R01AG032990, 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.

## Author contributions

Designed the study: A.M.Z., C.B.L., L.W.J.; supervised the study: A.M.Z., C.B.L., L.W.J., D.H., I.M.; performed experiments: J.T., J.D.L., I.M.; analyzed data: X.T., J.D.L.; writing, review, and editing: X.T., A.M.Z., J.T., J.D.L., I.M., D.H., C.B.L., L.W.J. All authors agree with the submission and have no conflicts of interest to disclose.

## 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.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-34787-4>.

**Correspondence** and requests for materials should be addressed to A.M.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023