1 **Title**: The restricted N-glycome of neurons is programmed during differentiation

2

3 **Authors:**

4 Katherine Kiwimagi[1], Maxence Noel[2], Murat Cetinbas[3], Ruslan I. Sadreyev[3], Lei Wang[1,#], Jordan

5 W. Smoller[4], Richard D. Cummings[2], Ron Weiss[1*], and Robert G. Mealer[5,6]*

6

7 **Affiliations:**

8 [1]Synthetic Biology Center, Department of Biomedical Engineering, Massachusetts Institute of

9 Technology, Cambridge, MA, USA.

10 [2]National Center for Functional Glycomics, Department of Surgery, Beth Israel Deaconess

11 Medical Center, Harvard Medical School, Boston, MA, USA.

12 [3]Department of Molecular Biology, Massachusetts General Hospital, Harvard Medical School,

13 Boston, MA, USA.

14 [4]Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine and

15 Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston,

16 MA, USA.

17 [5]Department of Psychiatry, School of Medicine, University of North Carolina at Chapel Hill,

18 Chapel Hill, NC, USA.

19 [6]Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill,

20 Chapel Hill, NC, USA.

21

22 **Running Title**: N-glycome of neurons is programmed during differentiation

23

24 **Keywords: Glycosylation, N-glycome, neurons, iPSCs**

25 ***Corresponding Authors: rweiss@mit.edu, rmealer@unc.edu**

26 **#Current address: Northeastern University, lei1.wang@northeastern.edu**

Abstract

The protein glycome of individual cell types in the brain is unexplored, despite the critical function of these modifications in development and disease. In aggregate, the most abundant asparagine (N-) linked glycans in the adult brain are high mannose structures, and specifically $Man_5GlcNAc_2$ (Man-5), which normally exits the ER for further processing in the Golgi. Mannose structures are uncommon in other organs and often overlooked or excluded in most studies. To understand cell-specific contributions to the unique brain N-glycome and its abundance of Man-5, we performed RNAseq and MALDI-MS TOF protein N-glycomics at several timepoints during differentiation of multiple cell types. To this end, homogeneous cultures of glutamatergic neurons, GABAergic neurons, and brain-specific endothelial cells were generated from monoclonal human inducible pluripotent stem cells (hiPSCs) through cellular reprogramming. Small molecule induction of stably integrated synthetic transcription units driving morphogen expression generated differentiated cells with distinct patterns mirroring intact tissue. Comparing uninduced hiPSCs for each cell type revealed identical transcriptomic and glycomic profiles before differentiation, with low quantities of Man-5. In differentiated glutamatergic and GABAergic neurons, the most abundant N-glycans became Man-5 and its immediate precursor Man-6, despite the presence of transcripts encoding enzymes for their subsequent modification. Differentiation to brain-specific endothelial cells showed an opposite effect, with the N-glycome displaying an abundance of complex N-glycans and terminal modifications of the late secretory pathway. These results confirm that the restricted N-glycome profile of brain is programmed into neuronal differentiation, with regulation independent of the transcriptome and under tight evolutionary constraint.

Introduction

N-glycosylation is an evolutionarily conserved biological pathway involving the enzymatic attachment of carbohydrate polymers to asparagine residues of proteins. N-glycosylation is critical in the development and function of the brain, as congenital disorders associated with the pathway most commonly present with severe neurologic phenotypes including seizures and intellectual disability[1]. Growing evidence is also revealing that common genetic variants in glycosylation genes are associated with more complex neuropsychiatric phenotypes including schizophrenia and Alzheimer's disease[2,3]. One such example is mutations in the manganese (Mn) transporter *SLC39A8*, as many glycosylation enzymes require Mn as a co-factor for activity. Severe mutations in *SLC39A8* result in a type II congenital disorder of glycosylation characterized by near total loss of circulating Mn and profound psychomotor impairment, epilepsy, and growth abnormalities[4,5]. On the opposite end of the allelic spectrum, a common missense variant (Ala391Thr - A391T, rs13107325) in *SLC39A8* is associated with hundreds of complex human phenotypes by GWAS[6], including decreased serum Mn[7], schizophrenia[8], intelligence[9], and several neuroimaging findings[10,11]. Both human carriers and mouse models of A391T have slightly reduced circulating Mn and abnormal serum glycosylation[12–14], and the A391T mouse model has altered protein glycosylation in the brain[15]. Interestingly, in the brain, *SLC39A8* is almost exclusively expressed in endothelial cells and absent in neurons[16].

Protein N-glycosylation is present in all domains of life but thought to be most complex in mammals[17]. All eukaryotic cells initiate N-glycosylation on the cytoplasmic side of the ER, where a dolichol-phosphate anchor is elongated with monosaccharides and then flipped into the lumen. This lipid-oligosaccharide complex is further extended in the ER prior to transfer onto asparagines (N-linked) of recently translated proteins entering the secretory pathway. The N-glycan precursor is then trimmed back by a series of α-glucosidases to generate the structure $Man_9GlcNAc_2$ (Man-9) and $Man_8GlcNAc_2$ (Man-8), which are common structures on

79   glycoproteins exiting the ER. Mannose residues are then sequentially cleaved in the cis-Golgi

80   by a series of α-mannosidases to generate $Man_5GlcNAc_2$ (Man-5). After high mannose (Man-5

81   through Man-9) structures, generation of complex N-glycans in all cells requires the action of the

82   N-acetylglucosaminyltransferase I (MGAT1), which adds a GlcNAc residue to Man-5. MGAT1 is

83   an essential enzyme, broadly expressed across all tissues with knock-out mice dying around

84   embryonic day 10 and showing notable impairments in neural development[18]. After MGAT1, N-

85   glycans can be modified by α-mannosidase II and hundreds of different glycosyltransferases

86   and related enzymes to generate a seemingly infinite array of glycan structures.

87

88   Most tissues and plasma display an abundance of complex N-glycans with relatively few high

89   mannose structures, generally less than <20%[19,20]. Some studies even exclude high mannose

90   structures from analysis, presumably as precursors for more complex and biologically relevant

91   glycans[21]. The brain N-glycome, however, is predominantly high mannose structures by bulk

92   (>60%), with Man-5 being the most abundant[22]. This observation is consistent across

93   mammals[22,23] as well as zebrafish[24]. We supposed one contributor to the unique N-glycome of

94   the brain was broad transcriptional downregulation of glycosylation genes compared to other

95   tissues[22]. In neurons, other groups have described unconventional secretory processing[25],

96   Golgi-independent trafficking[26] and activity-dependent Golgi satellites[27] as mechanisms which

97   contribute to the unique qualities and function of N-glycans in the brain. However, the

98   contribution of individual cell types as well as their developmental timeline responsible for the

99   distinct brain N-glycome is unknown.

100

101   To help fill these gaps in understanding, we generated human iPSCs with or without the

102   *SLC39A8* gene and subsequently differentiated them into glutamatergic neurons, GABAergic

103   neurons, and endothelial cells using inducible gene circuits, expecting only endothelial cells to

104   be affected by *SLC39A8* deletion. We confirmed specific differentiation through RNAseq

105   analysis and used immunofluorescence and qPCR as additional probes of cell state when

106   needed. MALDI-TOF MS revealed unique N-glycome profiles for both neuronal cell types and

107   endothelial cells, mirroring previous studies of intact tissue and primary cells[22,28]. Genomic

108   deletion of *SLC39A8* had minimal effect on the N-glycome of glutamatergic and GABAergic

109   neurons, whereas loss of this gene in endothelial cells where it is normally expressed caused a

110   dramatic reduction of complex N-glycans. These findings are the first detailed glycosylation

111   studies of individual human brain cell types and highlight the distinct and complex regulation of

112   this critical modification in the nervous system.

113

114   **Results**

115   **Morphogen expression via stably integrated inducible circuits drives targeted cellular**

116   **differentiation**

117   To study the effect of differentiation on the N-glycome of distinct cell types, gene driver circuits

118   were stably integrated into human induced pluripotent stem cells (hiPSCs) with a shared

119   isogenic background (PGP1)[29]. One endothelial and two neuronal cell types were targeted via

120   doxycycline (DOX) induced morphogen expression in gene driver circuits with previously

121   established differentiation protocols optimized for generation of homogeneous monocultures

122   free of support cells (**Fig. 1A**). The targeted neuronal cell types were based on inducible

123   expression of *NGN1* and *NGN2* (iNN) for glutamatergic neurons[30] and inducible expression of

124   *ASCL1* and *DLX2* (iAD) for GABAergic neurons[31]. In a similar fashion, brain-specific endothelial

125   cell induction was based on inducible expression of *ETV2* (iE)[32], though there is some

126   controversy as to whether *ETV2* expression alone is sufficient to achieve a brain specific

127   endothelial phenotype[33]. This controversy may be due to gaps in our knowledge on the effect of

128   multiple transcription factors and the interplay between them, as well experimental variables

129   such as varying growth media using different markers to define cell type. We therefore spent

5

130   additional time characterizing resultant endothelial cells to confirm their transcriptomic and

131   functional state. Circuits were stably integrated using Lenti or Piggybac into PGP1s followed by

132   monoclonal sorting as described in Methods. After stable integration and gated sorting,

133   monoclonal cell lines containing gene driver circuits were selected for differentiation into

134   homogeneous populations through DOX induction (+DOX), with subsequent analyses

135   performed on days 0, 4, and 14 (**Fig. 1B).**

136

137   In addition, a pair of otherwise identical cells lacking the expression of the Mn transporter

138   SLC39A8 (*SLC39A8-/-*) was created using CRISPR/Cas9 and verified using genomic PCR for

139   each gene driver set cell line (**Supp. Fig. 1**). As *SLC39A8* expression in the brain is limited

140   exclusively to endothelial cells[16], its deletion would be expected to have little to no effect on

141   neuronal cell types in homogeneous cultures. This resulted in 3 isogenic pairs (6 cell lines total)

142   to investigate N-glycome differences during differentiation. Transformation of hiPSCs to

143   differentiated monocultures was first confirmed by cellular morphology (**Fig. 1C**), followed by in

144   depth parallel analyses using bulk RNAseq transcriptomics and MALDI-TOF MS N-glycomics

145   (**Fig. 1D**).

146

147   Human iPSCs grown in culture without DOX induction (-DOX) maintained an iPSC phenotype,

148   with a patchy, sheet-like morphology and clusters of round cells seen on high magnification

149   (**Supp. Fig. 2**). After 14 days +DOX, both glutamatergic iNN and GABAergic iAD neurons

150   showed a less dense culture containing condensed somas and numerous neurite extensions. In

151   contrast, iE endothelial cells generated a dense monolayer of elliptical-appearing cells.

152   Immunofluorescence for the iNN cell line with antibodies for the pluripotency markers SOX2 and

153   OCT4 showed robust expression in the -DOX condition, which was then lost after 4 days +DOX

154   (**Supp. Fig. 3**). Expression of NeuN was present in the cytosol of -DOX iNN cells, consistent

155   with prior studies of undifferentiated cells (**Supp. Fig. 4**). After 4 days +DOX, expression of

156   NeuN localized to the nucleus, consistent with a neuronal phenotype[34]. Expression of the

157   neurofilament protein Tuj1 was absent in the -DOX iNN cells, but exhibited a strong signal in

158   neuronal processes with 4 days +DOX. These results demonstrate on a morphological and

159   protein level the conversion from iPSCs to differentiated cell types after induction.

160

161   **RNAseq confirms cell identity and differentiation state of homogeneous cultures**

162   Bulk RNAseq was performed from each cell line before and after differentiation aligned to the

163   human genome. Common cell markers for undifferentiated pluripotent cells (*DPPA4*, *NANOG*,

164   *SOX2*), neurons (*MAP2*, *GRIA2*, *NCAM1*), glutamatergic neurons (*VGLUT1*, *SATB2*, *SYN1*),

165   GABAergic neurons (*DLX1*, *GAD2*, *GAT1*), and endothelial cells (*FLI1*, *ERG*, *TAL1*) were

166   compared across different cell types and showed enriched expression in each of the intended

167   cell types (**Fig. 2**). Additional support for a brain-endothelial cell phenotype of ETV2 cell lines

168   included the presence of mature endothelial cell markers *CDH5*, *CLDN5*, VWF, *FLT1*, and *TIE1,*

169   and the reduction of epithelial markers *MUC1*, *EPCAM*, *CDH3*, and *CDH1* (**Supp. Table 1**).

170   RNAseq results for the induced morphogen followed the expected trend for iNN and iAD;

171   however, in iE cells, the uninduced population exhibited unexpected high expression of *ETV2* at

172   baseline (**Supp. Fig. 5**). To investigate this further, we performed qPCR using ETV2-specific

173   primers on multiple +/- DOX iE samples, which confirmed low expression of *ETV2* at baseline in

174   uninduced cells that increased in the presence of DOX (**Supp. Fig. 6**). As uninduced iE cell

175   lines maintain the iPSC morphology and expression of iPSC markers such as *NANOG* and

176   *SOX2,* we hypothesize that the confounding RNAseq results were a result of a lack in RNAseq

177   probe sensitivity or an off-target artifact of the iE genetic engineering cassette, though it did not

178   appear to affect differentiation or morphology.

179

180   A MDS plot comparing the global transcriptomic profile across all cell lines was consistent with

181   specific and distinct cell types (**Fig. 3**). All 6 undifferentiated iPSCs, independent of the gene

182   driver cassette, showed tight clustering of their transcriptomes. Following differentiation, iE lines

183   were closely associated and distinct from the other clusters, as were the iNN and iAD neuronal

184   cell lines. *SLC39A8* genotype appeared to have minimal effect on the transcriptomic profile of

185   the cell lines in comparison to +/-DOX and the gene driver expression cassette, though iNN KO

186   cells had minor separation from WT cells. These results are consistent with differentiation from

187   pluripotency to homogeneous and mature cell cultures at the RNA level.

188

189   **The N-glycome is dynamic and cell-type specific during differentiation**

190   We next performed MALDI-TOF MS glycomic analyzes of permethylated N-glycans prepared

191   from each cell line at day 0, 4, and 14 of differentiation. For each set of cells, N-glycan masses

192   were included if they had; 1- an isotopic appearance, 2 - a mass that corresponds to a possible

193   glycan composition, and 3 - the average signal to noise ratio was greater than 6 (S/N>6). This

194   resulted in the inclusion of 92, 80, 78, and 89 individual N-glycan masses for undifferentiated

195   iPSCs, iNN, iAD, and iE cultures, respectively. Upon differentiation, unique N-glycome profiles

196   emerged for each cell line, exhibiting distinctive overall profiles as well as exclusive peaks of

197   highest abundance (**Fig. 4, Table 1**).

198

199   **The N-glycome of human iPSCs resemble other less differentiated cells**

200   The N-glycome of undifferentiated iPSCs from every cell line was nearly identical (**Supp. Fig.**

201   **7**). High mannose structures composed ~75% of the signal, with Man-9 and Man-8 being the

202   most abundant structures, each independently representing ~21% of the total signal (**Table 1**,

203   **Supp. Fig. 8**). Of the complex N-glycans present (24% total), most contained a fucose (88%)

204   and many were sialylated (53%). The N-glycome of human embryonic stem cells exhibited a

205   similar pattern with high abundance of larger oligomannosidic structures including Man-8 and

206   Man-9[35]. This pattern is also observed in studies of other cultured cells including HEK-293s[36],

207   HeLa[37], and CHOs[38] – with the exception of CHO cells lacking the N-

208   acetylglucosaminyltransferase required for the final processing step of Man-5 to complex

209   glycans, MGAT1, known as Lec1 cells[38].

210

211   **The N-glycome of *NGN1/NGN2*-induced glutamatergic neurons (iNN) has abundant high**

212   **mannose structures including Man-5**

213   During *NGN1/NGN2* induction, iPSCs differentiating into glutamatergic lineage (iNN) exhibited a

214   progressive shift to structures of smaller glycan masses by day 4 which continued through day

215   14 (**Supp. Fig. 9**). High mannose structures composed 62.5% of the signal, with Man-5 being

216   the most abundant structure (24.5%) of the total signal, followed by Man-6 (15.6%) (**Table 1**,

217   **Supp. Fig. 10**). Of the complex N-glycans present, a similar amount contained a fucose

218   resideue compared to iPSCs, but far fewer were sialylated. Even the most abundant complex N-

219   glycan detected in iNN cells exhibited apparent less abundance than any of the high mannose

220   glycans, demonstrating the dramatic shift towards these glycans.

221

222   **The N-glycome of *ASCL1/DLX2*-induced GABAergic neurons (iAD) has abundant high**

223   **mannose glycans including Man-6 and Man-5**

224   After 4 days of *ASCL1/DLX2* induction, iPSCs differentiating into GABAergic lineage (iAD)

225   showed a slower shift towards the structures that would be most abundant by day 14 (**Supp.**

226   **Fig. 11**). At day 14, high mannose structures composed 58.5% of the signal, with Man-6 and

227   Man-5 being the most abundant structures at 16.7% and 14.1% of the total signal, respectively

228   (**Table 1**, **Supp. Fig. 12**). Of the complex N-glycans present, essentially all contained fucose, in

229   addition to several pauci-mannose structures with fucose, and far fewer complex glycans

230   contained sialic acid compared to both iNN and iPSCs. This pattern highlights a shift from

231   complex and sialylated N-glycans to high mannose N-glycans during GABAergic neuron

232   differentiation.

233

234 **The N-glycome of *ETV2*-induced endothelial cells (iE) displays robust sialylation and**

235 **glycan complexity**

236 At 4 days of *ETV2* induction, iPSCs differentiating into endothelial cells (iE) had begun to exhibit

237 a reduction of high mannose glycans and an increase of more complex glycans (**Supp. Fig. 13**).

238 At day 14, high mannose structures had decreased to 37.1 % of the signal, with Man-9

239 remained the most abundant structure of this class at 10.7% (**Table 1**, **Supp. Fig. 14**). Complex

240 N-glycans had increased substantially to 59.8%, the majority of which contained both fucose

241 and sialic acid. The complex biantennary glycan containing 2 galactoses, 1 fucose, and 1 sialic

242 acid (FA2G2S1), had increased from 2.0% in iPSCs to 11.4% and 18.5% by day 4 and day 14,

243 respectively, becoming the most abundant structure in iE cells (Supplementary Dataset 1). This

244 data suggests that during differentiation to endothelial cells from iPSCs, the N-glycome

245 transitions to more complex structures containing branches, fucose, and sialylation.

246

247 **The N-glycome of differentiated iPSCs tightly cluster and mirror intact tissue and primary**

248 **cells**

249 We next sought to quantitatively compare the glycomes of each cell line considering the entirety

250 of the glycomic profile. Glycans with an isotopic appearance and mass that corresponds to

251 known composition were selected at days 0, 4, and 14 of DOX induction for every sample.

252 Individual *m/z* values present in all samples were extracted and normalized within each sample,

253 followed by application of a Savitzky-Golay filter to reduce noise. A partial least squares

254 regression (PLSR) against days of DOX induction resulted in an optimal number of 4

255 components, and plotting the first two components revealed clear groupings (**Fig. 5**).

256 Resembling our RNAseq analysis, all 6 undifferentiated iPSCs showed tight clustering of their

257 glycomes, consistent with their overall similar profiles (**Supp. Fig. 7**). After 4 days of DOX

258 induction, each cell line had diverged from their parental iPSC glycome, though had not formed

259 clearly independent groups. By day 14, the neuronal cells formed a tight cluster, with only slight

260   differences seen between glutamatergic and GABAergic cell lines and minimal influence of

261   *SLC39A8* genotype. Incorporating N-glycomics summary data from our detailed study of adult

262   mouse cortex[22] into the model showed that our differentiated neuronal cell types had close

263   correlation to that of intact mouse brain tissue. Further, N-glycomics data from a human brain

264   sample from the same study showed similar clustering in the model although further separated

265   – however this single donor sample was from a commercial source with different processing.

266   Day 14 wild-type endothelial cells clustered with a previously published N-glycome of primary

267   human umbilical vein endothelial analyzed in a similar fashion (permethylation, MALDI-MS TOF)

268   and timepoint (14 days in culture)[28]. In contrast, the N-glycome of day 14 endothelial cells

269   lacking *SLC39A8* most closely associated with the group of day 4 induced cells. This finding

270   suggests that, despite expressing a transcriptome consistent with mature endothelial cells (**Fig.**

271   **3**), differentiated endothelial cells lacking *SLC39A8* have impaired N-glycome maturation

272   compared to wild-type cells.

273

274   **Combined transcriptome/analysis reveals unique secretory processing patterns of**

275   **neurons through the Golgi**

276   Given the distinct glycome profiles of iPSCs differentiating to neurons and endothelial cells, we

277   assessed the expression and enzymatic products of the core Golgi glycosyltransferases

278   involved in processing high mannose structures to generate complex N-glycans. Bottlenecks of

279   glycan synthesis are more easily appreciated by arranging structures along their linear path of

280   synthesis, akin to data from Bradberry and colleagues on fucosylation and synaptic vesicles[39].

281   Overlaying the subcellular localization and measured levels of each enzymatic transcripts

282   provides an additional level of information on differences in the secretory pathway between cell

283   lines (**Fig. 6A**). After leaving the ER, Man-9 is sequentially processed in the cis-Golgi by four

284   distinct class-I α-mannosidases to generate Man-5. In the medial-Golgi, this structure is

285   modified with a single GlcNAc by MGAT1, generating a glycan with one antenna (A1-Man-5).

11

286   A1-Man-5 is then cleaved by two interchangeable class-II α-mannosidases, generating the

287   pauci-mannose templates used by glycosyltransferases of the trans-Golgi to generate the

288   complex N-glycan structures common to most intact tissues.

289

290   In our prior work[22] and that of other groups using different analytic techniques[20,23,39], the N-

291   glycome of intact brain tissue shows an abundance of high mannose structures and a buildup of

292   Man-5, which is consistently the most abundant N-glycan detected. In comparison,

293   undifferentiated iPSCs show low levels of Man-5 and high levels of both early cis-Golgi

294   structures (Man-9) and complex N-glycans having completed Golgi processing (**Fig. 6B**). In

295   comparison, both iNN and iAD neurons exhibited a build-up of cis-Golgi products including Man-

296   5 and Man-6, respectively, mirroring the intact brain N-glycome. The N-glycome of iE cells show

297   lower relative levels of all high-mannose structures and a dramatic increase of complex N-

298   glycans, similar to primary HUVECs[28].

299

300   We and others have shown correlations between glyco-gene transcripts and their predicted

301   products in brain and cell lines[22,36]. Thus, one potential explanation for the accumulation of Man-

302   5 in neurons and the brain would be a lack of *MGAT1* expression, as is the case with Lec1 CHO

303   cells[38]. However, RNAseq results from mouse cortex as well as each homogeneous culture

304   confirmed comparable expression of *MGAT1* (**Fig. 6B**). Further, the expression of the four

305   class-I alpha-mannosidases proximal and at least one of two interchangeable class-II α-

306   mannosidases distal to *MGAT1* are expressed in all the cell lines to similar levels. These results

307   further highlight the distinct N-glycome maturation process inherent to neuronal differentiation,

308   which cannot be directly predicted by transcriptional analysis.

309

310   **Discussion**

311   In this study, we used complementary methods on distinct cell-types differentiated from a

312   shared hiPSC line to explore the cell-specific development of the brain N-glycome. We

313   demonstrate that the unique profile of the brain is driven by neuronal cell types, both excitatory

314   and inhibitory, and programmed into their differentiation. Further, we used this system to explore

315   cell-specific effects of deleting the schizophrenia risk-gene *SLC39A8*, which, in contrast with

316   most risk alleles, is expressed exclusively in endothelial cells.

317

318   The mechanisms regulating the distinct brain glycome and its function remain understudied[40],

319   particularly the connections and changes observed in disorders of the brain[41]. Given the

320   complexity of the brain compared to other tissues, in which N-glycan complexity is relatively well

321   distributed between high mannose, hybrid, and complex-type N-glycans in terms of abundance,

322   it could be predicted that the N-glycome would perhaps be the most elaborate. However, the

323   opposite is observed. Relatively few structures represent the majority of brain N-glycans, with a

324   predominance of high-mannose, particularly Man-5, which is often considered simple as an

325   intermediate precursor of more complex glycans with itself having unclear biological

326   significance. Hanus and colleagues demonstrated the presence of high mannose structures on

327   many critical proteins at the neuronal surface in culture[25]. They reported that blocking N-glycan

328   processing beyond high mannose structures with kifunensine, deoxymannojirimycin, and

329   swainsonine, all inhibitors of α-mannosidases, had no effect on dendritic arborization. In

330   contrast, more proximal inhibition of N-glycan synthesis with tunicamycin, which blocks the

331   initiation of N-glycan synthesis and high mannose generation on dolichol precursors, results in

332   impaired dendritic arborization. Recent work from the same group showed that high mannose

333   glycans are abundant at the neuronal surface in intact brains, and that changing their

334   abundance alters the electrophysiologic properties of related circuits[42]. These studies provide

335   support for a functional role of high mannose glycans at the surface of neurons and highlight the

336   need for understanding the accumulation of these structures on a cellular level.

13

337

338   Several groups have investigated regional and developmental changes to the brain N-glycome

339   across multiple mammal species. Lee and colleagues reported an in-depth study using LC-MS

340   analyzed 9 different brain regions in adult mice, as well several timepoints in mouse and human

341   prefrontal cortex[20]. They report a strong conservation of the overall N-glycome profile across

342   each sample, with small but significant changes noted in less abundant structures including

343   sialylated and fucosylated glycans[20]. Our group analyzed four brain regions (cortex,

344   hippocampus, striatum, and cerebellum) in mice using MALDI-TOF MS and described a

345   conserved N-glycome characterized by high mannose structures, as well as correlations

346   between gene expression with glyco-transcripts involved in bisection and fucosylation[22]. Studies

347   by Klarić and colleagues compared the same four brain regions in rat, macaque, chimpanzee,

348   and humans at two time points and using combinations of chromatography and mass

349   spectrometry, overlaid with gene expression data, and observed overall similar regional

350   patterns[23]. In each study, subtle but statistically significant changes are observed between

351   samples, and often used to draw conclusions regarding the function and evolution of the brain

352   glycome. For example the change from ~45.0% to ~50.0% abundance of complex N-glycans

353   from rat to human, or the reduction in α-2,3-linked sialic from ~0.7% to ~0.5% and a

354   corresponding increase in α-2,6-linked sialic acid[23]. Analytical differences likely exist between

355   studies, such as elution or ionization properties of different types of glycans in LC-MS vs

356   MALDI. Despite this, an overlooked aspect of the N-glycome in every species, tissue, and time-

357   point, is that high mannose glycans, and specifically Man-5, are always the most abundant

358   structures in brain tissue. We show here that this build-up of Man-5 and its immediate precursor

359   Man-6 is developmentally programmed into neuronal differentiation and should be considered

360   the distinguishing feature of the mammalian brain N-glycome in comparison to other cell types

361   and tissues, which commonly display much more complex N-glycans.

362

363   All eukaryotes share the same conserved core N-glycan processing steps and enzymes[17], but

364   different animals may differ in their overall diversity of N-glycans. For example, while both the

365   drosophila and mammalian brain N-glycome are dominated by high mannose structures[43], the

366   peripheral tissues in drosophila are also primarily high-mannose structures, in contrast to the

367   peripheral tissues of vertebrates, which display a diverse array of complex and modified N-

368   glycans[44]. It is thought that the expansion in complexity of peripheral N-glycans of vertebrates,

369   like other glycan classes, is driven by evolution, generating new carbohydrate structures at the

370   cell surface for adaptive advantages such as managing the increased number of cell-cell

371   interactions, regulating cell signaling, including interactions with microbes and pathogens[45–48].

372   The brain is separated from the periphery and pathogens by the blood-brain barrier, and thus

373   may not require such complex glycans to identify and battle invading organisms. It is however

374   the most complex organ in terms of cellular organization, cell-cell interactions, and cellular

375   diversity. Thus, the restricted N-glycome of neurons and the brain compared to other cell types

376   and tissues may result from not needing more complex N-glycans to function at some basal

377   level, or, in contrast, that such strong constraint on the synthesis of complex N-glycans exists

378   that relatively few are made. The conservation of the brain N-glycome across species, as well

379   as the profound neurological symptoms which are hallmark of glycosylation disorders, suggests

380   the latter is more likely.

381

382   A lack of MGAT1 expression would result in Man-5 accumulation, as is found in Lec1 cells[38],

383   though neurons express comparable levels compared to other cell types. *Mgat1*-null drosophila

384   have decreased lifespan and abnormal brain development characterized by fused lobes[49]. Mice

385   lacking *Mgat1* do not survive past mid-gestation (E9.5-E10.5), and show impaired neural tube

386   formation[18]. The human *MGAT1* gene is under considerable genetic constraint[50], and no known

387   human disorders of MGAT1-deficiency have been reported. Targeted inhibition of *Mgat1* in the

388   hippocampus of adult mice affected spine density and some electrophysiologic parameters

389   including evoked potentials, but there was no report of histological neuronal death or

390   phenotypes such as seizures in these animals[42]. As such, the role of high mannose and

391   complex N-glycans during both development and normal function in the adult brain necessitate

392   further study. Interestingly, all cells and tissues we have analyzed show a very low abundance

393   of A1-Man-5 (< 0.7%), an essential intermediate structure for the synthesis of all complex N-

394   glycans and the product of MGAT1. This suggests that the subsequent processing steps from

395   A1-Man-5 by MAN2A1 and MAN2A2 to generate the pauci-mannose structures, upon which all

396   complex N-glycans are built, is highly efficient when these enzymes can access their substrates.

397

398   An alternative explanation for the accumulation of Man-5 in neurons is the physical separation

399   of maturing glycans in the secretory pathway from MGAT1. This is supported by several lines of

400   evidence, including the existence of unique secretory processing and Golgi-independent

401   trafficking of secretory vesicles in neurons[25,26], as well as the weak correlation of

402   glycosyltransferase expression levels with the observed N-glycome[23]. Further, glycome

403   prediction tools based on glyco-gene expression levels, such as GlycoMaple, do not easily

404   predict the buildup of Man-5 in brain tissue[36]. As such, although the transcripts for enzymes

405   involved in generating complex N-glycans are expressed in neurons, their products many

406   seldom interact with their substrates, or may do so only with a limited population or under

407   specific conditions. Green and colleagues have described the existence of Golgi-satellites in

408   neurons which show activity dependent activation of N-glycans[27]. A study by Stanley and

409   colleagues identified an endogenous protein inhibitor of MGAT1, though this gene (*MGAT4D*)

410   appears to only be expressed in the testis[51]. Spatial separation of different N-glycan species in

411   neurons is also supported by enrichment of different glycan classes in certain subcellular

412   structures[39], as well as distinct patterns of N-glycan-binding lectins in the mammalian brain[52].

413   Lectins which bind high mannose glycans, such as ConA and GNL, show a diffuse pattern in the

16

414   brain, while lectins for fucosylated and bisected N-glycans, specifically AAL and PHA-E, display

415   enhanced signal in the synapse-rich molecular layer of the cerebellum.

416

417   Aside from the dramatic changes seen in congenital disorders of glycosylation and related

418   mouse models, most glycome changes in common human diseases are modest in size[41], and

419   may relate more to downstream consequences of having the condition rather than playing a role

420   in the etiology[53]. For example, our group compared the N-glycoproteome of 10 normal, 10

421   asymptomatic, and 10 symptomatic Alzheimer disease brain samples, and although some

422   subtle differences were noted between groups, the overall profiles are strikingly similar[54]. An

423   exception may be schizophrenia, where multiple glycosylation enzymes and related genes,

424   including *SLC39A8*, are genetically associated with risk of developing the disorder[2]. In this

425   study, we explored the cell-specific effect of *SLC39A8* deletion on the N-glycomes of neurons

426   and endothelial cells. In contrast to the majority of schizophrenia risk genes, which display

427   enriched expression in neurons[8], *SLC39A8* is expressed exclusively in brain endothelial cells[16].

428   Akin to human plasma glycome studies[12,14], we previously reported that mice harboring the

429   schizophrenia risk variant A391T in *Slc39a8* showed changes in glycosylation of proteins

430   expressed exclusively in neurons, suggesting a non-cell-autonomous effect of the schizophrenia

431   risk variant on protein glycosylation. In line with those results, we find that *SLC39A8* deletion in

432   neurons had minimal effect on their protein N-glycomes, whereas deletion in endothelial cells

433   caused a dramatic reduction in the branching of complex N-glycans. Although some studies

434   suggest that *SLC39A8* deletion or the A391T variant expressed in neuronal cell types can affect

435   their function and contribute to schizophrenia risk[55], the implications of these are unclear as

436   *SLC39A8* is not endogenously expressed in these cell types[16].

437

438   In summary, the results here emphasize both the unique N-glycome profile of neurons, which is

439   programmed into their cell-specific differentiation, and the necessity of studying disease-

440   relevant mutations in their appropriate cellular context. Future studies involving protein

441   glycosylation in the brain should consider other less abundant brain cell types, including

442   astrocytes, oligodendrocytes, and microglia, as well as other related pathways such as other

443   glycosylation pathways, e.g., O-GalNAc glycosylation. Further, co-cultures and co-differentiation

444   protocols will help understand the dynamic changes of the brain N-glycome and inform targeted

445   treatments for abnormal protein glycosylation in the brain.

446

447   **Author Contributions:**

448   KK was involved in initial conceptualization of the project, generated all cell lines, performed

449   differentiation experiments and cluster analyses for glycomic data, and wrote the manuscript

450   MN helped perform and optimize glycan assays and performed data analysis for glycomics

451   experiments

452   MC performed analysis of RNAseq data

453   RIS designed and supervised RNAseq analysis

454   LW designed and constructed integration vectors and generated polyclonal hiPSC lines used in

455   developing iE cells

456   JWS was involved in initial conceptualization of the project, experimental design, and data

457   analysis

458   RDC was involved in initial conceptualization of the project, experimental design, data analysis,

459   and oversaw all glycosylation analyses

460   RW was involved in initial conceptualization of the project, experimental design, data analysis,

461   oversaw all cell generation and validation experiments, and wrote the manuscript

462   RGM was involved in initial conceptualization of the project, performed all glycomic experiments

463   and data analyses, and wrote the manuscript

464

465   All authors contributed feedback and edits to the manuscript

18

466

467   **Author ORCID:**

468   KK - 0000-0002-5844-9412

469   MN - 0000-0003-2621-7374

470   LW - 0000-0001-5218-556X

471   JWS - 0000-0002-0381-6334

472   RDC - 0000-0002-8918-5034

473   RW - 0000-0003-0396-2443

474   RGM - 0000-0001-5963-2836

475

480

481   **Competing Interests:** The authors declare no competing interests.

482

483   **Methods**

484   **Generation of Isogenic Cell lines**

485   *Cell Maintenance.* Standard iPCS culture practices were used for all cell line maintenance. In

486   brief, Matrigel (VWR Cat # 354277) coated plates were used with daily mTeSR+ (Stem Cell

487   Technologies Cat #100-0276) media changes. Cells were aggregate or single-cell passaged,

488   depending on the downstream application, with Gentle Cell Dissociation Reagent (Stem Cell

489   Technologies Cat # 7174). Cells were gently thawed by rolling the cryovial between gloved

490   hands until fully liquid and then placed directly into newly prepared plates with media. Rock

491   inhibitor (StemCell Technologies Cat # 72304) was used in all media for thawing, passaging and

492   in any case where colony size was significantly small, less than ~10 cells. Cells were frozen

493   down in dissociated suspensions of ~1,000,000 cells per ml in mFreSR media (Stem Cell

494   Technologies Cat # 05854). Mr. Frosty (ThermoFisher Scientific Cat # 5100-0001) was used to

495   slowly cool cells during the freezing process. Cells were transferred to -140°C when fully

496   frozen. All cell lines were generated from the same previously published human iPSC line,

497   termed PGP1[29]. We confirmed that this cell line was derived from a male donor who was

498   homozygous for the major allele (C) at rs13107325, which codes for alanine (A) at amino acid

499   position 391 in *SLC39A8*.

500

501   *Plasmids*. Three base cell lines were initially created for this project, iNN (inducible circuit for

502   glutamatergic neuron differentiation via _NGN1_ and _NGN2_ expression[30]), iAD (inducible circuit for

503   GABAergic neuron differentiation via _ASCL1_ and _DLX12_[31]) and iE (inducible circuit for

504   endothelial cell differentiation via *ETV2*[32]). Addgene plasmids containing the Tetracycline-

505   dependent promoter (TRE-tight) driving respective transcription factors (#61471 (NGN1/NGN2)

506   or #97330 (DLX2) with #97329 (*ASCL1*)) were paired with an additional Lentivirus vector

507   containing the core promoter for human Elongation Factor-1 α (hEF1a1) driving constitutive

508   expression of reverse tetracycline-controlled transactivate (rtTA) to create the glutamatergic and

509   GABAergic circuits respectively (Fig. 1A). The piggyBac integration system was used to create

510   our ETV2 cell lines using a Golden Gate cloning strategy similar to previous work[56]. This

511   modular cloning strategy was expanded to allow for construction of two transcriptional units

512   (TUs) in a single piggyBac integration vector. For our PiggyBac transposon vector, the first TU

513   consisted of TRE3G mediated *ETV2-2A-YFP*, the second TU used hEF1a to drive rtTA3, and

514   the third TU applied hEF1 to drive the hygromycin resistance gene. To create cells lacking the

515   expression of *SLC39A8* (SLC39A8-KO), a similar modular cloning strategy was used to create

516   expression vectors for CRISPR/Cas9 and the U6 RNA polymerase III promoter (U6) driving

20

517   guide RNA expression. Four guide RNAs were designed and initially tested on wild type PGP1.

518   Only 1 combination gave successful modifications in the wild type PGP1, CHOPCHOP1 and

519   CHOPCHOP3. Therefore, these two guide RNAs were used to make cuts that remove a 156 bp

520   piece in both alleles in all three of our inducible cell lines (Supp. Fig. 1). Sequences for each

521   guide RNA's target sequence can be found in Supplementary material Figure 1b.

522

523   *Lentiviral Production.* Lentivirus was created using a 3rd generation expression system through

524   transient transfection of HEK293FT cells based on established protocols[57]. Fresh media was

525   placed on wells 24 hr after transfection, and this media was harvested 48 hr after transfection.

526   Virus-contained media was then filtered through a 0.45 µm polyethersulfone (PES) membrane.

527   Virus was not frozen but rather used the same day of collection.

528

529   *Lentiviral transduction.* A Matrigel-coated 24-well plate of PGP1 cells at 50% confluency were

530   incubated with filtered virus at 3 different concentrations, 1:4, 1:12 & 1:36 virus:media, in a total

531   volume of 400 µl plus 10 µg/ml Polybrene mTeSR+ media with Rock for each well. Virus-

532   containing media was allowed to sit on cells for 24 hours before being changed. The resulting

533   polyclonal populations were expanded and underwent 2 passages before further analysis.

534

535   PiggyBac Integration: The piggyBac transposase mediated integration strategy was utilized to

536   stably integrate our iE system[58]. In short, our *ETV2* PiggyBac transposon vector was co-

537   transfected with a consultatively expressed PiggyBac transposase at the ratio of 4:1 for stable

538   integration of our circuit. 48 hours after transfection, the stable integrated hiPSCs were selected

539   with hygromycin for 14 days. The resulting polyclonal population was then expanded for sorting.

540

541   *Single Cell Sorting, Expansion, and Selection.* After stable integration of genetic circuits creating

542   polyclonal populations, single cell sorts were performed to create monoclonal populations for

21

543   each cell type. For iNN and iAD cell lines, sorting was gated based only on morphology to

544   identify single cell populations, specifically three morphology gates were employed to isolate

545   single cells based on forward and side scatter area, forward scatter width and height and side

546   scatter width and height. Additionally, transiently transfected color controls were also used to

547   back gate showing clear isolation of living single cells. For iE cells, sorting of single cells

548   population was based on single cells displaying the top 10% of signal in the FITC channel to

549   detect eYFP fluorescence, which should correlate with ETV2 expression based on the construct

550   design harboring a 2A bicistronic peptide linker. Single cells from each line were sorted into 96-

551   well Matrigel coated plates containing mTeSR+ media with Rock. After 1 day, sort media was

552   replaced, and the single cells were allowed to expand for 5 days. After 5 days, media was

553   removed and replaced with media without Rock, and cells left for an additional 3-5 days until the

554   media color indicated decreased pH due to growth. Cells from one 96-well were moved to two

555   wells of a 24-well plate. One of these wells was then further expanded into a 6-well plate, while

556   DOX was added to the second well to assess for morphology changes over 2 days. If no

557   morphological change was noted, both samples were discarded. For cells, that did show

558   morphological positive samples, the morphological assay after DOX was repeated, and if again

559   positive, the -DOX sample was selected and froze down at -80°C. Over 100 colonies were

560   screened for each base cell line (iNN, iAD and iE), and of those that tested positive on the

561   morphological screen (3-10%), one line for each base cell was selected and kept in culture for

562   *SLC39A8*-knockout generation.

563

564   *SLC39A8-Knockout (KO) Generation.* Four guide RNA targeting human *SLC39A8* were

565   designed using the CHOPCHOP web application[59]. Two of these four (CHOPCHOP1 &

566   CHOPCHOP4) were designed to guide the system upstream of the Kozak sequence for

567   *SLC39A8* and the remaining two (CHOPCHOP2 & CHOPCHOP3) were designed to guide the

568   system downstream of the signaling sequence located in the first exon. All 4 logical combos

569   were tested through transient transfection into undifferentiated PGP1 cell lines in a 24-well

570   plate. Cells were grown for 2-3 days and then single cell sorted into Matrigel coated 96-well

571   plates with Rock containing media. However, instead of the DOX morphology screen, the

572   second split from the monoclonal sample was pelleted, and genomic DNA was extracted using

573   QuickExtract (Lucigen Cat # QE09050), and a PCR over the region of interest was then

574   conducted. During the initial screen only one of the four guide RNA pairs yielded a double

575   positive KO (ChopChop1 with ChopChop3). Generation of KOs in the remaining cell lines iNN,

576   iAD and iE used the top preforming pair of guide RNAs. One PCR product from a successful KO

577   was gel extracted using Monarch DNA Extraction Kit (NEB Cat # T1020) and set for

578   sequencing. The sequencing revealed the correct target region had been removed. One positive

579   cell line from each cell population was chosen, resulting in 3 isogenic cell pairs with and without

580   KO of *SLC39A8* (i.e. 6 cell lines total). Genomic PCR results from the final 6 cell lines confirmed

581   successful deletion of the region of interest in SLC39A8 as described in results. All 6 cell lines

582   further used in the study showed normal chromosomal reads as indicated by KaryoStat Assays

583   conducted by Life Technologies (Catalog # A52849).

584

585   *Transfection.* Lipofectamine Stem Transfection Reagent (ThermoFisher Scientific cat#

586   STEM00001) was used to transfect all plasmids in Matrigel coated 24-well plates as follows:

587   167 ng total DNA was prepared by using equal ratios of all plasmids prepped at 100 ng/µl

588   involved in a particular transfection. Of note, no transfection marker, was used in order to

589   reduce chances of off target effects. Optimem (ThermoFisher Scientific cat#31985062) was

590   added to bring the total volume of each sample to 25 µl. A master mix of 1 µl Lipofectamine to

591   24 µl of Optimem was created, and 25 µl of this master mix was then added to each sample,

592   creating 50 µl total volume in each sample. The complex was incubated at room temperature for

593   10 min before being added to the cells. Cells were then reverse transfected as single cell

594   suspensions of 90,000 cells/ml and were added to Matrigel coated 24-well plates immediately

595   before the transfection complexes were added to the cells to increase the surface area the

596   transfection complexes could access on the cells. Plates were then incubated at 37°C overnight

597   to allow cells to adhere, and the media was changed the following day.

598

599   *DOX Induction.* 5 mg of Doxycycline Hyclate (DOX) (Sigma-Aldrich Cat# D9891) was

600   resuspended in nuclease free water H_2O to bring to make a 10 mg/ml stock, which was then

601   aliquoted and stored at –20C. DOX was added to the media of newly seeded 6-well plates at a

602   5 μM concentration. Media was changed every 2-3 days for iNN and iAD cells, and daily for iE

603   cells due to there more rapid proliferation.

604

605   *Harvesting Cells.* Undifferentiated iPSCs and iE cells were incubated at 37°C in Gentle Cell

606   Dissociation Reagent (Stem Cell Technologies CAT#100-0485) for 5 minutes. Gentle Cell was

607   aspirated and 1 ml PBS (VWR Cat#21-040-CV) was added, and cells were manually removed

608   using a cell scraper, transfer to a 15 mL conical tube, and then pelleted at 300 G for 3 minutes.

609   Supernatant PBS was carefully aspirated and pellets were flash frozen in liquid nitrogen. For

610   iNN and iAD neurons, the Gentle Cell incubation was skipped as the neurons were easily

611   dissociated from the plate. The media and cells were transferred to a 15 mL conical tube, and 1

612   mL was used to wash any additional cells off the plate and transferred to the conical tube, and

613   the cells were pelleted at 300 G for 3 minutes. The supernatant was carefully removed, and

614   pellets were flash frozen in liquid nitrogen. Of note, each culture and timepoint was grown with 6

615   technical replicates in a 6-well plate. Five technical replicates were pooled, pelleted and flash

616   frozen for glycome analyses. One technical replicate was pelleted and frozen for RNA-seq

617   analysis (note: we did not process the 4-day RNA-seq samples due to time constraints).

618

619 *Immunofluorescence.* iNN cells at 4 days +/- DOX were stained using standard protocols. In

620 brief, live cells were gently washed with ice cold PBS 2x before being covered with 4%

621 paraformaldehyde (Electron Microscopy Sciences Cat #157-4) for 20 min at room temperature

622 for fixation. Fixed cells were then washed with PBS 3x and incubated with permeability/blocking

623 buffer (10% donkey serum [Sigma Cat# D9663], 0. 2% Triton X-100 [MP Biomedicals cat

624 #04807426] in PBS) for 1 hr at room temperature. Primary antibodies were incubated for 1hr at

625 room temperature in the dark, washed with PBS 3x (5 min per wash), incubated with secondary

626 antibody, when indicated for 1 hr at room temperature in the dark, washed with PBS 4x (5 min

627 per wash), and then imaged using a Leica TCS SP5 II confocal microscope. The following

628 antibody combinations/dilutions in permeability/blocking buffer were used: Anti-Sox2 conjugated

629 to AlexaFluor647 (BD Biosciences Cat # 562139) at 1:200; Anti-Oct3/4 conjugated to

630 AlexaFluor555 (BD Biosciences cat # 560306) at 1:100; Anti-Fox3/NeuN (BioLegend cat #

631 834502) at 1:1,000, donkey anti-mouse AlexaFluor546 (Life Technologies cat # A10036) at

632 1:2,000; anti-Tuj1 (Neuromics cat # CH23005) at 1:100, goat anti-chicken AlexaFlour633 (Life

633 Technologies cat # A21103) at 1:2,000. Cultures were co-stained for either Oct4 and Sox2 or

634 NeuN and Tuj1 together.

635

636 *RNA isolation and Analysis.* RNA from snap frozen cell pellets was extracted using the Monarch

637 Total RNA Miniprep Kit (Cat # T2010S). RNAseq analysis was performed by the MGH NextGen

638 Sequencing Core as previously described[15,22]. In brief RNA-seq libraries were prepared from

639 total RNA using polyA selection followed by the NEBNext Ultra II Directional RNA Library Prep

640 Kit protocol (New England Biolabs, E7760S). Sequencing was performed on Illumina HiSeq

641 2500 instrument resulting in approximately 30 million 50 bp reads per sample. Sequencing

642 reads were mapped in a splice-aware fashion to the human reference transcriptome using

643 STAR [60]. Read counts over transcripts were calculated using HTSeq based on the Ensembl

25

644   annotation for GRCh38 assembly and presented as Transcripts Per Million (TPM)[61]. The RNA-

645   Seq statistical analysis (and related figures) was performed by using EdgeR package[62].

646

647   *MALDI-TOF N-Glycomic Analysis.* All samples were processed as previously described[22], with

648   slight modifications as noted below. Cell pellets were lysed in 500 µL ice-cold buffer (50 mM

649   Tris, 150 mM NaCl, 1.0% w/v Triton-X-100, pH 7.6) with protease inhibitor (Roche

650   #46931320019), followed by brief dissociation using a hand-held motorized pestle (Kimble

651   #749540) and 2 brief pulses of sonication for 10 seconds with a microtip (Qsonica Q700).

652   Volume was adjusted to 1 mL with additional lysis buffer, and protein concentrations were

653   measured using the Pierce BCA Protein Assay Kit (ThermoFisher Scientific #23255).

654   Glycoproteins from 1 mg of protein lysate were then dialyzed, lyophilized, reduced, dialyzed,

655   lyophilized, trypsinized, purified, eluted and lyophilized as previously described[22]. N-glycans

656   were released from lyophilized glycopeptides after resuspension in 200 µL of 50 mM ammonium

657   bicarbonate and incubated with 3 µL PNGase F (New England Biolabs, #P0704) at 37 °C for 4

658   h, then overnight (12–16 h) with an additional 3 µL of enzyme at 37 °C. Peptides were removed

659   by a preconditioned C18 Sep-Pak columns (200 mg Waters, #WAT054945) glycans were eluted

660   with 6 mL of 5% acetic acid, placed in a speed vacuum to remove organic solvents then and

661   lyophilized prior to permethylation as previously described[22]. Permethylated glycans were

662   resuspended in 25 µL of 75% methanol and spotted in a 1:1 ratio with DHB matrix on an MTP

663   384 polished steel target plate (Bruker Daltonics #8280781). MALDI-TOF MS N-glycomics data

664   was acquired from a Bruker Ultraflex II instrument using FlexControl Software in the reflective

665   positive mode with a mass/charge (m/z) range of 1,000–5,000 kD. Twenty independent

666   captures (representing 1,000 shots each) were obtained from each sample and averaged to

667   create the final combined spectra file. Data was exported in msd format using FlexAnalysis

668   Software for subsequent annotation. For each set of cells, N-glycan masses were included if

669   they had; 1- an isotopic appearance based on m/z, 2 - a mass that corresponds to a possible N-

26

670   glycan composition, and 3 - the average signal to noise ratio was greater than 6 (S/N > 6). This

671   resulted in the inclusion of 92, 80, 78, and 89 individual N-glycan masses for undifferentiated

672   iPSCs, iNN, iAD, and iE cultures, respectively. The relative abundance of each glycan was

673   calculated as the signal intensity for each isotopic peak divided by the summed signal intensity

674   for all measured glycans within a spectrum. N-glycans were grouped into different categories

675   based on shared components, such as monosaccharide composition, antennarity, etc., and the

676   summed abundance of each category was compared. All glycan structures are presented

677   according to the Symbol Nomenclature for Glycans (SNFG) guidelines[63] and were drawn using

678   GlycoGlyph[64]. A Partial Least Square Regression Analysis (PLSR) was performed using the

679   normalized abundance from every peak that passed the criteria described above (isotopic,

680   known composition, S/N >6) from any of the cell lines, resulting in 118 individual m/z values for

681   comparison. If a peak was not originally annotated in the spectra for a cell line (not meeting the

682   above-described criteria for the initial analysis) it was subsequently flagged in the other samples

683   for the PLSR analysis such that there were no gaps in the peak data across samples. After

684   normalization to relative abundance (signal intensity for each peak divided by the summed

685   signal intensity for all measured peaks), the data was processed through a Savitzky-Golay filter

686   to reduce noise before being regressed against the data corresponding day of DOX induction

687   (0, 4 and 14). An optimal number of 4 components was found by minimizing the mean squared

688   error of the cross validation between the predicted day and the actual day of DOX induction.

689   The first 2 components were plotted against each other to assess clustering. Published data for

690   permethylated N-glycans measured using similar techniques by MALDI-MS TOF from human

691   cortex[22], mouse cortex[22], and primary endothelial cells[28] were processed using the same filters

692   and components and plotted for comparison.

693

694   *Data Availability.* The data generated during this study are included in this published article and

695   its supplementary information files, and available from the corresponding author on reasonable

696   request. Raw MS glycomics and RNAseq data will be made available in public databases upon

697   publication of this manuscript.

698   **References:**

700   1. Freeze, H.H., Eklund, E.A., Ng, B.G., and Patterson, M.C. (2015). Neurological aspects of
701   human glycosylation disorders. Annu. Rev. Neurosci. *38*, 105–125.
702   https://doi.org/10.1146/annurev-neuro-071714-034019.

703   2. Mealer, R.G., Williams, S.E., Daly, M.J., Scolnick, E.M., Cummings, R.D., and Smoller, J.W.
704   (2020). Glycobiology and schizophrenia: a biological hypothesis emerging from genomic
705   research. Mol. Psychiatry *25*, 3129–3139. https://doi.org/10.1038/s41380-020-0753-1.

706   3. Tang, X., Tena, J., Di Lucente, J., Maezawa, I., Harvey, D.J., Jin, L.-W., Lebrilla, C.B., and
707   Zivkovic, A.M. (2023). Transcriptomic and glycomic analyses highlight pathway-specific
708   glycosylation alterations unique to Alzheimer's disease. Sci. Rep. *13*, 7816.
709   https://doi.org/10.1038/s41598-023-34787-4.

710   4. Park, J.H., Hogrebe, M., Grüneberg, M., DuChesne, I., von der Heiden, A.L., Reunert, J.,
711   Schlingmann, K.P., Boycott, K.M., Beaulieu, C.L., Mhanni, A.A., et al. (2015). SLC39A8
712   Deficiency: A Disorder of Manganese Transport and Glycosylation. Am. J. Hum. Genet. *97*,
713   894–903. https://doi.org/10.1016/j.ajhg.2015.11.003.

714   5. Boycott, K.M., Beaulieu, C.L., Kernohan, K.D., Gebril, O.H., Mhanni, A., Chudley, A.E.,
715   Redl, D., Qin, W., Hampson, S., Küry, S., et al. (2015). Autosomal-Recessive Intellectual
716   Disability with Cerebellar Atrophy Syndrome Caused by Mutation of the Manganese and
717   Zinc Transporter Gene SLC39A8. Am. J. Hum. Genet. *97*, 886–893.
718   https://doi.org/10.1016/j.ajhg.2015.11.002.

719   6. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C.,
720   McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS
721   Catalog of published genome-wide association studies, targeted arrays and summary
722   statistics 2019. Nucleic Acids Res. *47*, D1005–D1012. https://doi.org/10.1093/nar/gky1120.

723   7. Ng, E., Lind, P.M., Lindgren, C., Ingelsson, E., Mahajan, A., Morris, A., and Lind, L. (2015).
724   Genome-wide association study of toxic metals and trace elements reveals novel
725   associations. Hum. Mol. Genet. *24*, 4739–4745. https://doi.org/10.1093/hmg/ddv190.

726   8. Trubetskoy, V., Pardiñas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B.,
727   Bryois, J., Chen, C.-Y., Dennison, C.A., Hall, L.S., et al. (2022). Mapping genomic loci
728   implicates genes and synaptic biology in schizophrenia. Nature *604*, 502–508.
729   https://doi.org/10.1038/s41586-022-04434-5.

730   9. Hill, W.D., Marioni, R.E., Maghzian, O., Ritchie, S.J., Hagenaars, S.P., McIntosh, A.M.,
731   Gale, C.R., Davies, G., and Deary, I.J. (2019). A combined analysis of genetically correlated
732   traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Mol.
733   Psychiatry *24*, 169–181. https://doi.org/10.1038/s41380-017-0001-5.

734   10. Luo, Q., Chen, Q., Wang, W., Desrivières, S., Quinlan, E.B., Jia, T., Macare, C., Robert,
735   G.H., Cui, J., Guedj, M., et al. (2019). Association of a Schizophrenia-Risk Nonsynonymous

736   Variant With Putamen Volume in Adolescents: A Voxelwise and Genome-Wide Association
737   Study. JAMA Psychiatry *76*, 435. https://doi.org/10.1001/jamapsychiatry.2018.4126.

738   11. Hermann, E.R., Chambers, E., Davis, D.N., Montgomery, M.R., Lin, D., and Chowanadisai,
739   W. (2021). Brain Magnetic Resonance Imaging Phenome-Wide Association Study With
740   Metal Transporter Gene SLC39A8. Front. Genet. *12*, 647946.
741   https://doi.org/10.3389/fgene.2021.647946.

742   12. Mealer, R.G., Jenkins, B.G., Chen, C.-Y., Daly, M.J., Ge, T., Lehoux, S., Marquardt, T.,
743   Palmer, C.D., Park, J.H., Parsons, P.J., et al. (2020). The schizophrenia risk locus in
744   SLC39A8 alters brain metal transport and plasma glycosylation. Sci. Rep. *10*, 13162.
745   https://doi.org/10.1038/s41598-020-70108-9.

746   13. Nakata, T., Creasey, E.A., Kadoki, M., Lin, H., Selig, M.K., Yao, J., Lefkovith, A., Daly, M.J.,
747   Graham, D.B., and Xavier, R.J. (2020). A missense variant in SLC39A8 confers risk for
748   Crohn's disease by disrupting manganese homeostasis and intestinal barrier integrity. Proc.
749   Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.2014742117.

750   14. Sunuwar, L., Frkatović, A., Sharapov, S., Wang, Q., Neu, H.M., Wu, X., Haritunians, T.,
751   Wan, F., Michel, S., Wu, S., et al. (2020). Pleiotropic ZIP8 A391T implicates abnormal
752   manganese homeostasis in complex human disease. JCI Insight *5*.
753   https://doi.org/10.1172/jci.insight.140978.

754   15. Mealer, R.G., Williams, S.E., Noel, M., Yang, B., D'Souza, A.K., Nakata, T., Graham, D.B.,
755   Creasey, E.A., Cetinbas, M., Sadreyev, R.I., et al. (2022). The schizophrenia-associated
756   variant in SLC39A8 alters protein glycosylation in the mouse brain. Mol. Psychiatry *27*,
757   1405–1415. https://doi.org/10.1038/s41380-022-01490-1.

758   16. Saunders, A., Macosko, E.Z., Wysoker, A., Goldman, M., Krienen, F.M., de Rivera, H., Bien,
759   E., Baum, M., Bortolin, L., Wang, S., et al. (2018). Molecular Diversity and Specializations
760   among the Cells of the Adult Mouse Brain. Cell *174*, 1015-1030.e16.
761   https://doi.org/10.1016/j.cell.2018.07.028.

762   17. Stanley, P., Moremen, K.W., Lewis, N.E., Taniguchi, N., and Aebi, M. (2022). N-Glycans. In
763   Essentials of Glycobiology, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M.
764   Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, et al., eds. (Cold Spring
765   Harbor Laboratory Press).

766   18. Ioffe, E., and Stanley, P. (1994). Mice lacking N-acetylglucosaminyltransferase I activity die
767   at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates.
768   Proc. Natl. Acad. Sci. *91*, 728–732. https://doi.org/10.1073/pnas.91.2.728.

769   19. Wallace, E.N., West, C.A., McDowell, C.T., Lu, X., Bruner, E., Mehta, A.S., Aoki-Kinoshita,
770   K.F., Angel, P.M., and Drake, R.R. (2024). An N-glycome tissue atlas of 15 human normal
771   and cancer tissue types determined by MALDI-imaging mass spectrometry. Sci. Rep. *14*,
772   489. https://doi.org/10.1038/s41598-023-50957-w.

773   20. Lee, J., Ha, S., Kim, M., Kim, S.-W., Yun, J., Ozcan, S., Hwang, H., Ji, I.J., Yin, D., Webster,
774   M.J., et al. (2020). Spatial and temporal diversity of glycome expression in mammalian
775   brain. Proc. Natl. Acad. Sci. U. S. A. *117*, 28743–28753.
776   https://doi.org/10.1073/pnas.2014207117.

21. Narimatsu, Y., Joshi, H.J., Nason, R., Van Coillie, J., Karlsson, R., Sun, L., Ye, Z., Chen, Y.-H., Schjoldager, K.T., Steentoft, C., et al. (2019). An Atlas of Human Glycosylation Pathways Enables Display of the Human Glycome by Gene Engineered Cells. Mol. Cell *75*, 394-407.e5. https://doi.org/10.1016/j.molcel.2019.05.017.

22. Williams, S.E., Noel, M., Lehoux, S., Cetinbas, M., Xavier, R.J., Sadreyev, R.I., Scolnick, E.M., Smoller, J.W., Cummings, R.D., and Mealer, R.G. (2022). Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. Nat. Commun. *13*, 275. https://doi.org/10.1038/s41467-021-27781-9.

23. Klarić, T.S., Gudelj, I., Santpere, G., Novokmet, M., Vučković, F., Ma, S., Doll, H.M., Risgaard, R., Bathla, S., Karger, A., et al. (2023). Human-specific features and developmental dynamics of the brain N-glycome. Sci. Adv. *9*, eadg2615. https://doi.org/10.1126/sciadv.adg2615.

24. Yamakawa, N., Vanbeselaere, J., Chang, L.-Y., Yu, S.-Y., Ducrocq, L., Harduin-Lepers, A., Kurata, J., Aoki-Kinoshita, K.F., Sato, C., Khoo, K.-H., et al. (2018). Systems glycomics of adult zebrafish identifies organ-specific sialylation and glycosylation patterns. Nat. Commun. *9*, 4647. https://doi.org/10.1038/s41467-018-06950-3.

25. Hanus, C., Geptin, H., Tushev, G., Garg, S., Alvarez-Castelao, B., Sambandan, S., Kochen, L., Hafner, A.-S., Langer, J.D., and Schuman, E.M. (2016). Unconventional secretory processing diversifies neuronal ion channel properties. eLife *5*, e20609. https://doi.org/10.7554/eLife.20609.

26. Bowen, A.B., Bourke, A.M., Hiester, B.G., Hanus, C., and Kennedy, M.J. (2017). Golgi-independent secretory trafficking through recycling endosomes in neuronal dendrites and spines. eLife *6*, e27362. https://doi.org/10.7554/eLife.27362.

27. Govind, A.P., Jeyifous, O., Russell, T.A., Yi, Z., Weigel, A.V., Ramaprasad, A., Newell, L., Ramos, W., Valbuena, F.M., Casler, J.C., et al. (2021). Activity-dependent Golgi satellite formation in dendrites reshapes the neuronal surface glycoproteome. eLife *10*, e68910. https://doi.org/10.7554/eLife.68910.

28. Jaillet, C., Morelle, W., Slomianny, M.-C., Paget, V., Tarlet, G., Buard, V., Selbonne, S., Caffin, F., Rannou, E., Martinez, P., et al. (2017). Radiation-induced changes in the glycome of endothelial cells with functional consequences. Sci. Rep. *7*, 5290. https://doi.org/10.1038/s41598-017-05563-y.

29. Lee, J.-H., Park, I.-H., Gao, Y., Li, J.B., Li, Z., Daley, G.Q., Zhang, K., and Church, G.M. (2009). A Robust Approach to Identifying Tissue-Specific Gene Expression Regulatory Variants Using Personalized Human Induced Pluripotent Stem Cells. PLoS Genet. *5*, e1000718. https://doi.org/10.1371/journal.pgen.1000718.

30. Busskamp, V., Lewis, N.E., Guye, P., Ng, A.H., Shipman, S.L., Byrne, S.M., Sanjana, N.E., Murn, J., Li, Y., Li, S., et al. (2014). Rapid neurogenesis through transcriptional activation in human stem cells. Mol. Syst. Biol. *10*, 760. https://doi.org/10.15252/msb.20145508.

31. Barretto, N., Zhang, H., Powell, S.K., Fernando, M.B., Zhang, S., Flaherty, E.K., Ho, S.-M., Slesinger, P.A., Duan, J., and Brennand, K.J. (2020). ASCL1- and DLX2-induced

817   GABAergic neurons from hiPSC-derived NPCs. J. Neurosci. Methods *334*, 108548.
818   https://doi.org/10.1016/j.jneumeth.2019.108548.

819   32. Morita, R., Suzuki, M., Kasahara, H., Shimizu, N., Shichita, T., Sekiya, T., Kimura, A.,
820   Sasaki, K., Yasukawa, H., and Yoshimura, A. (2015). ETS transcription factor ETV2 directly
821   converts human fibroblasts into functional endothelial cells. Proc. Natl. Acad. Sci. *112*, 160–
822   165. https://doi.org/10.1073/pnas.1413234112.

823   33. Lu, T.M., Houghton, S., Magdeldin, T., Durán, J.G.B., Minotti, A.P., Snead, A., Sproul, A.,
824   Nguyen, D.-H.T., Xiang, J., Fine, H.A., et al. (2021). Pluripotent stem cell-derived epithelium
825   misidentified as brain microvascular endothelium requires ETS factors to acquire vascular
826   fate. Proc. Natl. Acad. Sci. *118*, e2016950118. https://doi.org/10.1073/pnas.2016950118.

827   34. Mullen, R.J., Buck, C.R., and Smith, A.M. (1992). NeuN, a neuronal specific nuclear protein
828   in vertebratesxs. Development *116*, 201–211. https://doi.org/10.1242/dev.116.1.201.

829   35. Satomaa, T., Heiskanen, A., Mikkola, M., Olsson, C., Blomqvist, M., Tiittanen, M., Jaatinen,
830   T., Aitio, O., Olonen, A., Helin, J., et al. (2009). The N-glycome of human embryonic stem
831   cells. BMC Cell Biol. *10*, 42. https://doi.org/10.1186/1471-2121-10-42.

832   36. Huang, Y.-F., Aoki, K., Akase, S., Ishihara, M., Liu, Y.-S., Yang, G., Kizuka, Y., Mizumoto,
833   S., Tiemeyer, M., Gao, X.-D., et al. (2021). Global mapping of glycosylation pathways in
834   human-derived cells. Dev. Cell *56*, 1195-1209.e7.
835   https://doi.org/10.1016/j.devcel.2021.02.023.

836   37. Gao, W., Jiang, Y., Zhang, Z., Zhang, Y., Liu, Y., Zhou, Y., and Liu, X. (2017). A facile
837   method for cellular N-glycomic profiling by matrix-assisted laser desorption/ionization mass
838   spectrometry. RSC Adv. *7*, 35687–35693. https://doi.org/10.1039/C7RA06071H.

839   38. North, S.J., Huang, H.-H., Sundaram, S., Jang-Lee, J., Etienne, A.T., Trollope, A., Chalabi,
840   S., Dell, A., Stanley, P., and Haslam, S.M. (2010). Glycomics Profiling of Chinese Hamster
841   Ovary Cell Glycosylation Mutants Reveals N-Glycans of a Novel Size and Complexity. J.
842   Biol. Chem. *285*, 5759–5775. https://doi.org/10.1074/jbc.M109.068353.

843   39. Bradberry, M.M., Peters-Clarke, T.M., Shishkova, E., Chapman, E.R., and Coon, J.J.
844   (2023). N-glycoproteomics of brain synapses and synaptic vesicles. Cell Rep. *42*, 112368.
845   https://doi.org/10.1016/j.celrep.2023.112368.

846   40. Klarić, T.S., and Lauc, G. (2022). The dynamic brain N-glycome. Glycoconj. J. *39*, 443–471.
847   https://doi.org/10.1007/s10719-022-10055-x.

848   41. Pradeep, P., Kang, H., and Lee, B. (2023). Glycosylation and behavioral symptoms in
849   neurological disorders. Transl. Psychiatry *13*, 154. https://doi.org/10.1038/s41398-023-
850   02446-x.

851   42. Zhang, C.-L., Moutoussamy, C., Tuffery, M., Varangot, A., Piskorowski, R., and Hanus, C.
852   (2024). Core-N-glycans are atypically abundant at the neuronal surface and regulate
853   glutamate receptor signaling. Preprint, https://doi.org/10.1101/2024.03.25.586577
854   https://doi.org/10.1101/2024.03.25.586577.

855 43. Koles, K., Lim, J.-M., Aoki, K., Porterfield, M., Tiemeyer, M., Wells, L., and Panin, V. (2007).
856   Identification of N-Glycosylated Proteins from the Central Nervous System of Drosophila
857   Melanogaster. Glycobiology *17*, 1388–1403. https://doi.org/10.1093/glycob/cwm097.

858 44. Katoh, T., and Tiemeyer, M. (2013). The N's and O's of Drosophila glycoprotein
859   glycobiology. Glycoconj. J. *30*, 57–66. https://doi.org/10.1007/s10719-012-9442-x.

860 45. Gagneux, P., Panin, V., Hennet, T., Aebi, M., and Varki, A. (2022). Evolution of Glycan
861   Diversity. In Essentials of Glycobiology, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley,
862   G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, et al., eds.
863   (Cold Spring Harbor Laboratory Press).

864 46. Springer, S.A., and Gagneux, P. (2013). Glycan evolution in response to collaboration,
865   conflict, and constraint. J. Biol. Chem. *288*, 6904–6911.
866   https://doi.org/10.1074/jbc.R112.424523.

867 47. Schjoldager, K.T., Narimatsu, Y., Joshi, H.J., and Clausen, H. (2020). Global view of human
868   protein glycosylation pathways and functions. Nat. Rev. Mol. Cell Biol. *21*, 729–749.
869   https://doi.org/10.1038/s41580-020-00294-x.

870 48. Cummings, R.D (2024). Evolution and Diversity of Glycomolecules from Unicellular
871   Organisms to Humans. BioCosmos New Perspect. Orig. Evol. Life *4*, 1–35.
872   https://doi.org/10.2478/biocosmos-2024-0001.

873 49. Sarkar, M., Leventis, P.A., Silvescu, C.I., Reinhold, V.N., Schachter, H., and Boulianne, G.L.
874   (2006). Null mutations in Drosophila N-acetylglucosaminyltransferase I produce defects in
875   locomotion and a reduced life span. J. Biol. Chem. *281*, 12776–12785.
876   https://doi.org/10.1074/jbc.M512769200.

877 50. Chen, S., Francioli, L.C., Goodrich, J.K., Collins, R.L., Kanai, M., Wang, Q., Alföldi, J.,
878   Watts, N.A., Vittal, C., Gauthier, L.D., et al. (2024). A genomic mutational constraint map
879   using variation in 76,156 human genomes. Nature *625*, 92–100.
880   https://doi.org/10.1038/s41586-023-06045-0.

881 51. Huang, H.-H., Hassinen, A., Sundaram, S., Spiess, A.-N., Kellokumpu, S., and Stanley, P.
882   (2015). GnT1IP-L specifically inhibits MGAT1 in the Golgi via its luminal domain. eLife *4*,
883   e08916. https://doi.org/10.7554/eLife.08916.

884 52. Noel, M., Cummings, R.D., and Mealer, R.G. (2023). N-glycans show distinct spatial
885   distribution in mouse brain. Glycobiology *33*, 935–942.
886   https://doi.org/10.1093/glycob/cwad077.

887 53. Williams, S.E., Mealer, R.G., Scolnick, E.M., Smoller, J.W., and Cummings, R.D. (2020).
888   Aberrant glycosylation in schizophrenia: a review of 25 years of post-mortem brain studies.
889   Mol. Psychiatry *25*, 3198–3207. https://doi.org/10.1038/s41380-020-0761-1.

890 54. Suttapitugsakul, S., Stavenhagen, K., Donskaya, S., Bennett, D.A., Mealer, R.G., Seyfried,
891   N.T., and Cummings, R.D. (2022). Glycoproteomics Landscape of Asymptomatic and
892   Symptomatic Human Alzheimer's Disease Brain. Mol. Cell. Proteomics *21*, 100433.
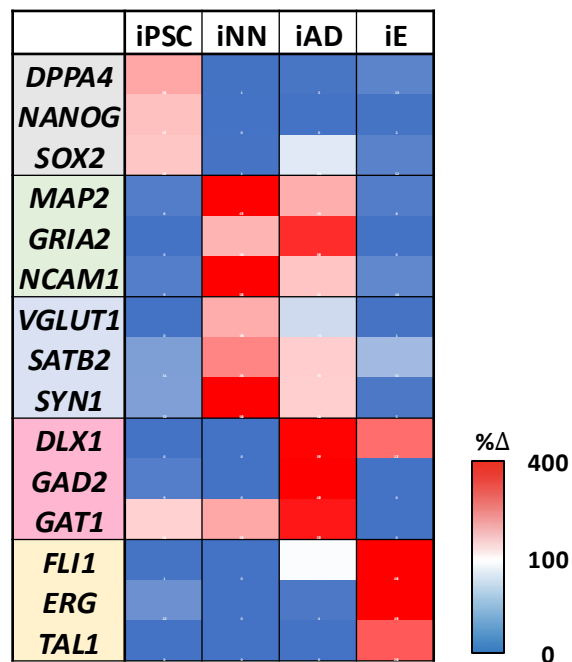893   https://doi.org/10.1016/j.mcpro.2022.100433.

894   55. Tseng, W.C., Reinhart, V., Lanz, T.A., Weber, M.L., Pang, J., Le, K.X.V., Bell, R.D.,
895   O'Donnell, P., and Buhl, D.L. (2021). Schizophrenia-associated SLC39A8 polymorphism is
896   a loss-of-function allele altering glutamate receptor and innate immune signaling. Transl.
897   Psychiatry *11*, 136. https://doi.org/10.1038/s41398-021-01262-5.

898   56. Duportet, X., Wroblewska, L., Guye, P., Li, Y., Eyquem, J., Rieders, J., Rimchala, T., Batt,
899   G., and Weiss, R. (2014). A platform for rapid prototyping of synthetic gene networks in
900   mammalian cells. Nucleic Acids Res. *42*, 13440–13451.
901   https://doi.org/10.1093/nar/gku1082.

902   57. Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998).
903   A Third-Generation Lentivirus Vector with a Conditional Packaging System. J. Virol. *72*,
904   8463–8471. https://doi.org/10.1128/JVI.72.11.8463-8471.1998.

905   58. Zhao, S., Jiang, E., Chen, S., Gu, Y., Shangguan, A.J., Lv, T., Luo, L., and Yu, Z. (2016).
906   PiggyBac transposon vectors: the tools of the human gene encoding. Transl. Lung Cancer
907   Res. *5*, 120–125. https://doi.org/10.3978/j.issn.2218-6751.2016.01.05.

908   59. Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., and Valen, E.
909   (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing.
910   Nucleic Acids Res. *47*, W171–W174. https://doi.org/10.1093/nar/gkz365.

911   60. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P.,
912   Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner.
913   Bioinforma. Oxf. Engl. *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

914   61. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-
915   throughput sequencing data. Bioinforma. Oxf. Engl. *31*, 166–169.
916   https://doi.org/10.1093/bioinformatics/btu638.

917   62. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package
918   for differential expression analysis of digital gene expression data. Bioinforma. Oxf. Engl.
919   *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

920   63. Neelamegham, S., Aoki-Kinoshita, K., Bolton, E., Frank, M., Lisacek, F., Lütteke, T.,
921   O'Boyle, N., Packer, N.H., Stanley, P., Toukach, P., et al. (2019). Updates to the Symbol
922   Nomenclature for Glycans guidelines. Glycobiology *29*, 620–624.
923   https://doi.org/10.1093/glycob/cwz045.

924   64. Mehta, A.Y., and Cummings, R.D. (2020). GlycoGlyph: a glycan visualizing, drawing and
925   naming application. Bioinforma. Oxf. Engl. *36*, 3613–3614.
926   https://doi.org/10.1093/bioinformatics/btaa190.

927
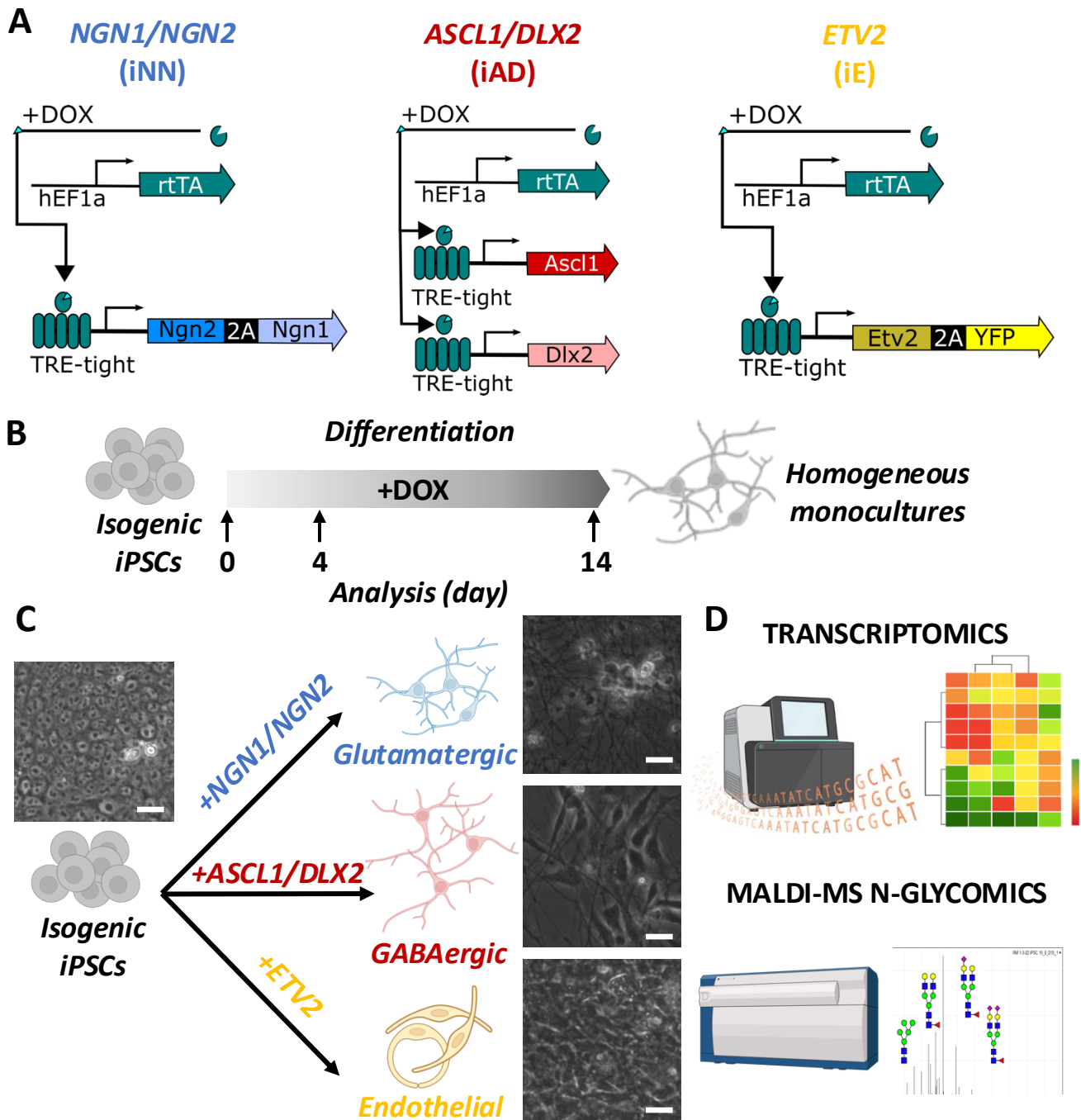
Figure 1. Defining the N-glycome of single human cell types. A) Schematic of differentiation and analysis of isogenic human iPSCs used to generate homogeneous monocultures of distinct cell types following doxycycline (+DOX) induction. B) Gene circuit designs for differentiation of glutamatergic neurons (iNN), GABAergic neurons (iAD), and endothelial cells (iE) based on established protocols. C) Schematic and light microscopy images showing differentiation of iPSCs into homogeneous monocultures with morphology reflecting distinct cell types. Scale bar = 20 μm. Uncropped images are presented in supplemental material. D) Analytic methods including RNAseq and MALDI-TOF MS N-glycomics applied at different time points confirm cell-specific transcriptome and glycome profiles.

	iPSC	iNN	iAD	iE
DPPA4				
NANOG				
SOX2				
MAP2				
GRIA2				
NCAM1				
VGLUT1				
SATB2				
SYN1				
DLX1				
GAD2				
GAT1				
FLI1				
ERG				
TAL1				

%Δ
400
100
0

Figure 2. Transcriptomic profiles are consistent with programmed cell type and differentiation state. Relative expression of common cell-type markers are shown for iPSCs (gray), all neurons (green), glutamatergic neurons (blue), GABAergic neurons (red), and endothelial cells (yellow). Data is presented as percent change (%Δ) of individual cultures compared to average expression across all cultures. Heat map scaled from 0-400 % based on average expression across cultures.
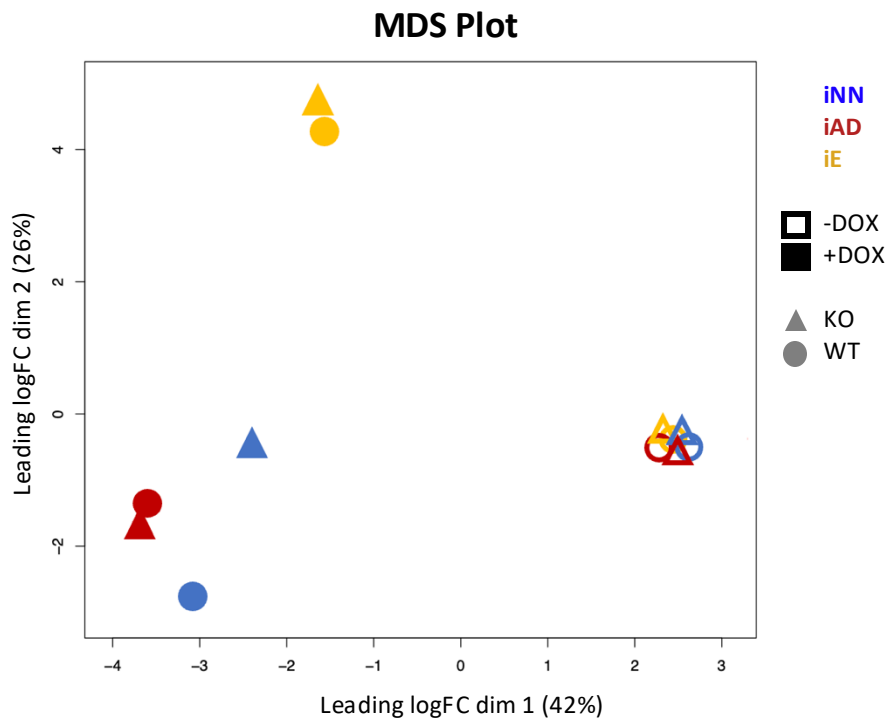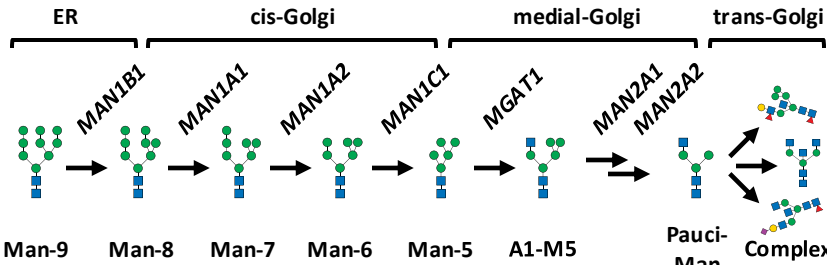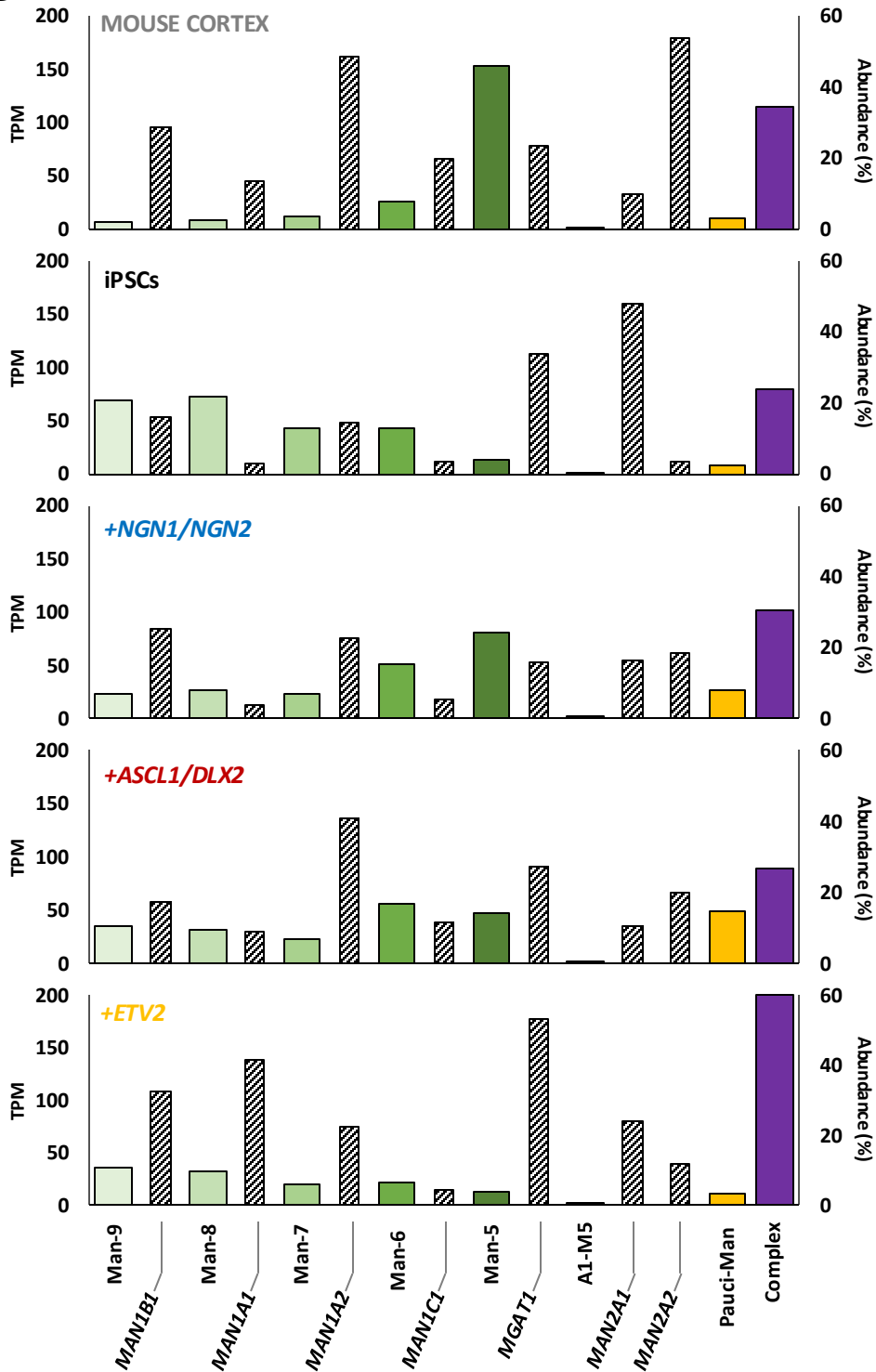
Figure 3. Transcriptomic profiles of distinct cell types form tight clusters. MDS plot of transcriptomic profiles from all cultures, highlighting the tight clustering of each undifferentiated iPSCs line (-DOX, open shapes) compared to induced cell lines (+DOX, filled shapes) expressing distinct gene drivers for iNN-neurons (blue), iAD-neurons (red) and iE-endothelial cells (yellow). Data is shown from two different genetic backgrounds, wild-type (circles) and *SLC39A8 -/-* (triangles) cells.
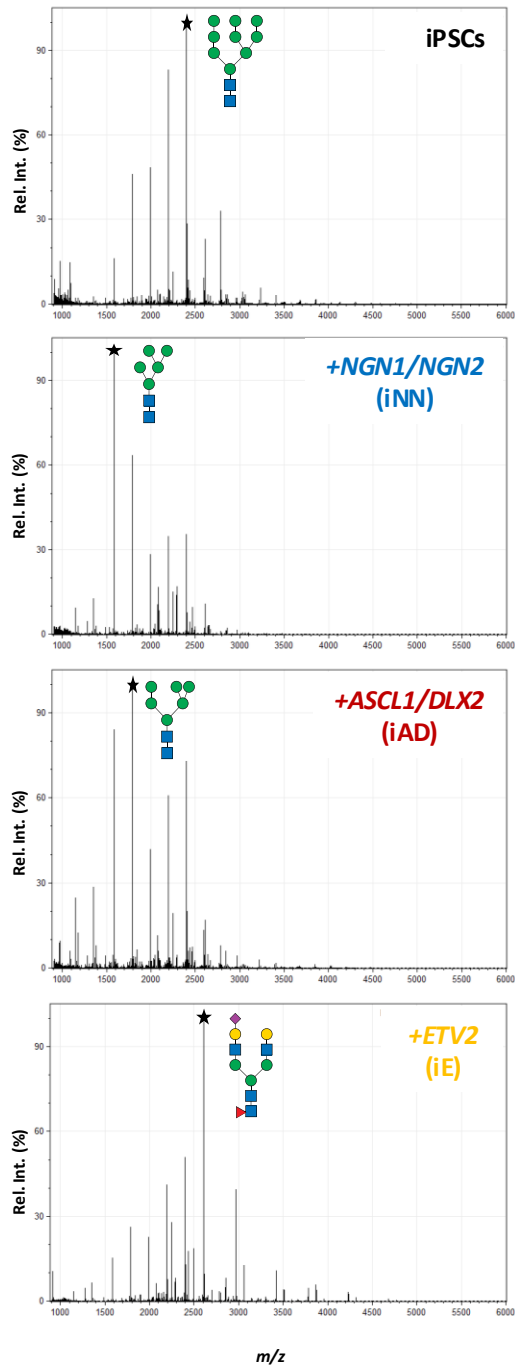
Figure 4. N-glycome profiles are unique between cell-type and resemble intact tissue.
MALDI-MS TOF analysis of permethylated N-glycans isolated from homogeneous wild-type cultures of undifferentiated and differentiated iPSCs. The corresponding structure for the most intense peak (★) in each culture is illustrated, including Man-9 (iPSCs), Man-5 (iNN), Man-6 (iAD), and FA2G2S1 (iE).
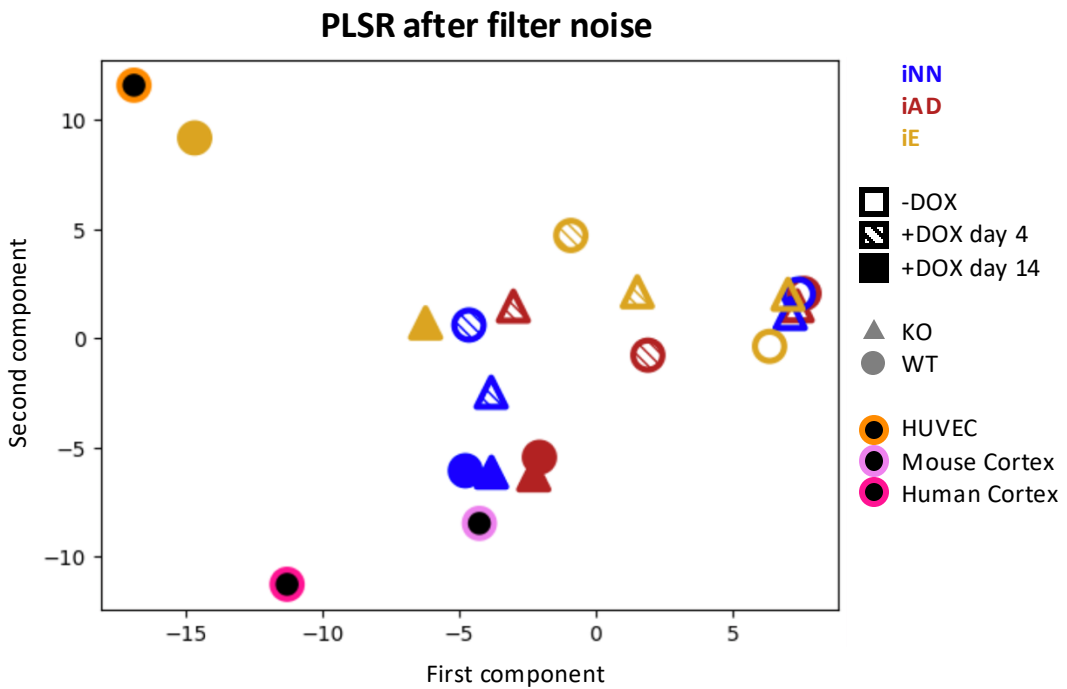
Figure 5. N-glycome profiles are unique between cell-type and resemble intact tissue.
Plot of glycomic profiles from all cultures using Partial Least Squares Regression (PLSR) analysis, highlighting the tight clustering of each undifferentiated iPSCs line (-DOX, open shapes) compared to induced cell lines (+DOX, filled shapes, 4-day hash fill, 14-day solid fill) expressing distinct gene drivers for iNN-neurons (blue), iAD-neurons (red) and iE-endothelial cells (yellow). Data shown from two different genetic backgrounds, wild-type (circles) and *SLC39A8 -/-* (triangles) cells. Reference N-glycomics data from primary human endothelial cells (HUVECs) (black circle with orange halo) and mouse cortex (black circle with light pink halo show tight clustering with related 14-day wild-type samples. Human cortex (black circle with fuchsia halo) is located in similar location though further south-west compared to derived neurons and mouse cortex cluster.

A

ER — cis-Golgi — medial-Golgi — trans-Golgi

MAN1B1 | MAN1A1 | MAN1A2 | MAN1C1 | MGAT1 | MAN2A1 MAN2A2

Man-9 → Man-8 → Man-7 → Man-6 → Man-5 → A1-M5 → Pauci-Man / Complex

B

MOUSE CORTEX

iPSCs

+NGN1/NGN2

+ASCL1/DLX2

+ETV2

Figure 6. N-glycans of the cis-Golgi accumulate in neurons despite the presence of *MGAT1*. A) Schematic of N-glycan synthesis from the ER through the Golgi, highlighting glycan types, glycogenes, and subcellular compartments. **B)** The relative abundance of individual N-glycans (%) from mouse cortex and differentiated neuronal cell types show a similar accumulation of the late cis-Golgi structures Man-6 and Man-5, while undifferentiated iPSCs and differentiated endothelial-like cells show an enrichment of structures from the ER and trans-Golgi, respectively. Glycogene transcript levels (diagonal stripes) for each step are shown as transcripts per million (TPM). Accumulation of the late cis-Golgi structures Man-6 and Man-5 in mouse cortex and differentiated neuronal cell types cannot be accounted for exclusively by an abundance of *MAN1C1* or a lack of *MGAT1*, as each express levels comparable to most tissues.

	iPSC	iNN	iAD	iE
Pauci-mannose	2.3	7.9	14.6	3.1
High-mannose	73.5	62.5	58.5	37.1
Complex	24.1	30.4	26.9	59.8
Sialic Acid	12.8	6.5	7.3	39.6
Fucose	21.2	27.0	29.6	52.6
Man-5	4.2	24.5	14.1	3.9
Man-6	12.8	15.6	16.7	6.7
Man-7	13.1	7.0	7.0	5.8
Man-8	21.7	7.8	9.5	9.8
Man-9	21.0	7.0	10.3	10.7

Table 1. Glycan abundances vary between human cell types after 14 days of induced differentiation. Data is shown as the sum of the percent abundance for each category or glycan, normalized within the total N-glycome profile.