Cell type mapping of mild malformations of cortical development with oligodendroglial hyperplasia in epilepsy using single-nucleus multiomics

Isabella C. Galvão[1,2], Manuela Lemoine[1,2], Ludmyla Kandratavicius[2,3], Clarissa L. Yasuda[2,4], Marina K. M. Alvim[2,4], Enrico Ghizoni[2,4], Ingmar Blümcke[5], Fernando Cendes[2,4], Fabio Rogerio[2,3], Iscia Lopes-Cendes[1,2], Diogo F. T. Veiga[1,2,*]

[1]Department of Medical Genetics and Genomic Medicine, School of Medical Sciences, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil
[2]The Brazilian Institute of Neuroscience and Neurotechnology (BRAINN), Campinas, São Paulo, Brazil
[3]Department of Pathology, School of Medical Sciences, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil
[4]Department of Neurology, School of Medical Sciences, University of Campinas (UNICAMP), Campinas, São Paulo, Brazil
[5]Department of Neuropathology, Partner of the European Reference Network (ERN) EpiCARE, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[*]Corresponding author: dfveiga@unicamp.br

ABSTRACT

Objective

Mild malformations of cortical development with oligodendroglial hyperplasia in epilepsy (MOGHE) are brain lesions associated with focal epilepsy and characterized by increased oligodendroglial density, heterotopic neurons, and hypomyelination in the white matter. While previous studies have implicated somatic mutations in the *SLC35A2* gene, the cellular and molecular mechanisms underlying MOGHE pathogenesis remain elusive. To address this gap, this study aimed to systematically characterize the cell type composition and molecular alterations of MOGHE lesions at cellular resolution using single-nucleus multiomics profiling.

Methods

We performed single-nucleus multiomics sequencing to obtain paired gene expression and chromatin accessibility profiles of > 31,000 nuclei from gray matter and white matter regions of MOGHE lesions, and compared the results with publicly available neurotypical control datasets.

Results

The analysis of gray and white matter regions from two MOGHE patients revealed significant cellular composition alterations, including an oligodendrocyte expansion and heterotopic neurons within the subcortical white matter. We identified a distinct population of MOGHE-associated oligodendrocytes characterized by expressing genes related to immune response, myelination disruption, and epilepsy-related pathways. These oligodendrocytes shared a common transcriptional signature with oligodendrocytes in other neurological conditions involving white matter abnormalities. Further analysis of heterotopic neurons in MOGHE revealed the upregulation of genes associated with neuronal migration and the Wnt signaling pathway, suggesting a mechanism underlying their atypical localization.

Significance

This high-resolution cell type mapping of MOGHE lesions in clinical samples unveils neuronal and glial populations affected by the disease, and provides novel insights into the pathophysiological mechanisms of MOGHE.

Keywords

MOGHE; malformations of cortical development; epilepsy; single-nucleus RNA sequencing; single-nucleus ATAC sequencing; gene expression; oligodendrocytes

Key Points

- We provide a multimodal cellular atlas of the human cortical and subcortical regions affected in MOGHE
- MOGHE-associated oligodendrocytes exhibit immune response and myelination dysfunction
- Neuronal migration and Wnt signaling are upregulated in heterotopic neurons

69 **INTRODUCTION**

70 Malformations of cortical development (MCDs) encompass a varied group of disorders that arise
71 due to disruptions in the development of the cerebral cortex. MCDs can affect individuals across
72 all age groups, although symptoms typically begin in early childhood and extend to young
73 adulthood and may include epilepsy, developmental delays, and intellectual impairments[1]. MCDs
74 typically manifest as focal epilepsy and include focal cortical dysplasia (FCD) as well as mild
75 malformations of cortical development with oligodendroglial hyperplasia in epilepsy (MOGHE),
76 a neuropathological entity featuring abnormalities predominantly in the white matter, while
77 maintaining normal cortical lamination[2]. This latter characteristic distinguishes MOGHE from
78 FCD, which is characterized by abnormal cortical architecture associated or not with abnormal cell
79 types[3]. MOGHE was proposed as a distinct pathological entity in 2017[2], and was recently included
80 in the International League Against Epilepsy (ILAE) classification[4]. The increase in
81 oligodendroglia and the appearance of heterotopic neurons in the white matter, together with the
82 blurring of the gray-white matter junction, are the neuropathological hallmarks of this condition.
83 The affected white matter tissue presents with proliferation of oligodendroglial cells expressing
84 Olig2 and irregular areas of hypomyelination due to reduced myelin protein expression. Clinically,
85 MOGHE manifests as focal epilepsy and treatment might require surgical resection of the affected
86 region. MOGHE lesions occur most often in the frontal lobe and have early seizure onset, typically
87 in young children, but may also occur in the temporal lobe, and symptoms may manifest later even
88 in adulthood[5,6].

89 Genetic investigations have found that MOGHE patients frequently harbor somatic brain variants
90 in the *SLC35A2* gene, which encodes a UDP-galactose transporter[7]. One study reported that 45%
91 (9/20) of MOGHE cases presented truncating or frameshift variants in *SLC35A2*[7], while another
92 screening identified *SLC35A2* variants in 47.1% (16/35) of cases[8]. In agreement, the loss of
93 *Slc35a2* has been shown to recapitulate the MOGHE-like phenotype in conditional knockout
94 mouse models[9–11]. Other genes may also play a role, as demonstrated by a case identified with a
95 variant in the *NPRL2* gene[8]. However, the cellular and molecular mechanisms contributing to
96 MOGHE pathogenesis have not been systematically studied.

97 Single-nucleus RNA sequencing (snRNA-seq) and single-nucleus assay for transposase accessible
98 chromatin sequencing (snATAC-seq) have been applied to dissect the cellular diversity of brain

99 tissues in neurological conditions, enabling the discovery of dysregulated cell states involved in
100 diseases including focal MCDs[12–16]. For instance, we previously applied paired snRNA-seq and
101 snATAC-seq to identify cell populations involved in FCD type II pathology, which allowed us to
102 uncover vulnerable neuronal subtypes and activated microglial states in these cortical
103 malformations[15].

104 In this study, we employed multiomics single-nucleus profiling to systematically analyze the
105 cellular landscapes of both gray matter and subcortical white matter of MOGHE tissues, and
106 compared the results with publicly available control datasets. As a result, we characterized
107 perturbed oligodendroglial and neuronal cell populations associated with the disease. Our findings
108 provide novel insights into the pathophysiological mechanisms of MOGHE.

109

110 **METHODS**

111 **Clinical samples and neuropathological diagnosis**

112 Fresh brain samples were collected from two patients who underwent surgery for drug-resistant
113 focal epilepsy at the Hospital de Clínicas, University of Campinas. All procedures were approved
114 by the Research Ethics Board of our Institution (CAAE: 12112913.3.0000.5404), and written
115 informed consent was obtained from patients or their legal guardians before surgery.

116

117 Representative surgical specimens were either formalin-fixed and paraffin-embedded (FFPE) or
118 snap-frozen in liquid nitrogen and stored at −80°C. FFPE samples were submitted to routine
119 diagnostics in serial 4 µm-sections stained with hematoxylin and eosin (H&E) and submitted to
120 immunohistochemical reactions. For the latter protocol, the sections were exposed to antibodies
121 against Olig2 (glial marker; dilution 1:250, polyclonal, Millipore Sigma, cat#AB9610, Darmstadt,
122 Germany), and MAP2 (neuronal marker; 1:1,000, clone M13, Thermo Fisher, cat#13-1500,
123 Waltham, MA, USA), overnight at room temperature. Then, a detection solution containing the
124 secondary antibody and peroxidase (AdvanceTMHRP®, Dako, cat#K4068, Glostrup, Denmark;
125 or EnvisionTM Flex+, Dako, cat#K8002, Glostrup, Denmark) was added for 30 min at 37°C. 3,3-
126 diaminobenzidine (DAB) was used as a chromogenic substrate and counterstaining was performed
127 with hematoxylin. Negative controls (without primary antibody) were run concurrently with all

128 immunohistochemical reactions. Digitized images were obtained using a CS2 Aperio ScanScope
129 scanner (Aperio Technologies, Vista, CA, USA).
130
131 Samples were neuropathologically classified as MOGHE according to the latest guidelines of the
132 ILAE[4]. Specifically, the specimens showed an increase in oligodendroglial cells (> 2200 Olig2-
133 positive cells/mm^2) and an excess of heterotopic neurons in the white matter (>30 MAP2-positive
134 neurons/mm^2). No cortical cytoarchitectural changes were observed. After histopathological
135 confirmation of MOGHE, a mirror-frozen section of the tissue used for FFPE analysis was
136 dissected using a surgical scalp in a Petri dish on dry ice, allowing separation of the GM and WM.
137 Subsequently, tissue samples were submitted to molecular analyses as described below.

138

139 **Whole-exome sequencing and somatic mutations**

140 Genomic DNA was extracted from MOGHE-WM tissue using the PureLink Genomic DNA Mini
141 Kit (cat# K182001, Invitrogen) following the recommended protocol. Exome capture libraries
142 were generated using the SureSelect Human All Exon v8 kit (Agilent). Paired-end sequencing was
143 performed at 101 base pairs on each side of the DNA fragment on the Illumina NovaSeq 6000
144 platform. Deep sequencing obtained an average of 489 M reads per sample, with a mean depth
145 exome coverage of 530x.

146 Reads were aligned to the hg38 human reference genome using bwa mem 0.7.17[17]. Bam files were
147 sorted and duplicates were removed using Picard tools v.2.18.2. Indel realignment and base quality
148 score recalibration were performed with GATK[18]. Variants were called using the GATK haplotype
149 caller v4.0.5.1[18] and annotated using SnpEff v.4.3[19]. Somatic mutations including single-
150 nucleotide variants and indels were detected using Mutect2 (GATK v.4.0.5.1) with the 1000
151 Genomes panel of normals as controls.

152

153 **Nuclei isolation and multiome library construction**

154 Nuclei were isolated using the Chromium Nuclei Isolation kit (10X Genomics) according to the
155 manufacturer's protocol (CG000505, revision A), which is compatible with the Single Cell
156 Multiome ATAC + Gene Expression assay. Briefly, frozen tissue samples were dissociated with
157 a pestle in lysis buffer and passed through a nuclei isolation column. The isolated nuclei were then
158 resuspended in a debris removal buffer, followed by washes to clear any remaining debris.
159 Recovered nuclei were stained with Trypan blue and counted with a hemocytometer. Nuclei
160 integrity was evaluated under a microscope at 40X or 60X magnification, and nuclei were
161 considered viable if they were roundish and with an intact membrane. We then performed library
162 construction, which included open chromatin transposition and droplet (GEM) formation using a
163 Chromium Controller (10X Genomics). Libraries were prepared using the Multiome ATAC +
164 Gene Expression kit (10X Genomics) following the manufacturer's protocol (CG000338, Revision
165 E). The quality of the purified ATAC and RNA libraries was assessed using TapeStation (Agilent
166 Technologies).

167

168 **Sequencing and raw data processing**

169 Multiome libraries were sequenced on the Illumina NovaSeq 6000 platform according to the
170 sequencing depth and read length guidelines specified in the Multiome kit protocol (CG000338).
171 Demultiplexing, genome alignment, gene quantification, and peak accessibility analyses for single
172 nuclei were conducted using the Cell Ranger ARC pipeline. Specifically, fastq files were generated
173 with cellranger mkfastq, followed by processing with cellranger-arc count v.2.0.2 using the human
174 genome reference GRCh38 to produce count matrices for RNA and ATAC modalities.
175 Downstream data processing for both snRNA-seq and snATAC-seq were performed using Seurat
176 v.5.0.1[20] and Signac v.1.12.0[21] in R v.4.1.2.

177

178 **snRNA-seq analysis**

179 Ambient RNA decontamination was carried out using DecontX[22] as implemented in the celda[23]
180 v.1.10.0 R package. Briefly, RNA counts for each sample were processed with the decontX
181 function, utilizing the raw droplet matrix generated by cellranger count as background. DecontX

182 then employs a Bayesian approach to estimate and remove contamination in individual cells.
183 Nuclei were filtered using Seurat based on the following criteria: 1000 < UMI counts < 25,000,
184 nFeatures > 400, and mitochondrial read percentage (percent.mt) < 15. Doublet identification was
185 conducted for each sample using scDblFinder[24] v.1.8.0.43.

186 Standard Seurat processing and normalization steps were applied in each sample, including
187 NormalizeData, FindVariableFeatures, ScaleData, RunPCA, RunUMAP, FindNeighbors, and
188 FindClusters. The filtered counts matrix was log-normalized with regression to correct for
189 mitochondrial gene percentage. To correct batch effects across samples, we recomputed
190 NormalizeData, FindVariableFeatures and RunPCA on the merged object, and Harmony[25] v.1.2
191 was used to integrate PCA projections with automatic hyperparameter optimization and default
192 settings. Harmony's algorithm projects cells into a shared space where they cluster based on cell
193 type rather than sample-specific attributes like sequencing batch. Dimensionality reduction was
194 performed using RunUMAP with the first 40 principal components identified using the elbow plot.
195 Clustering was executed with FindNeighbors and FindClusters functions employing the smart
196 local moving (SLM) algorithm, with a resolution set to 1.2.

197 Cell type markers were obtained by applying the Wilcox rank-sum test to RNA-normalized data,
198 with a log2 fold-change threshold of 0.25 and an adjusted P-value of less than 0.05. Differentially
199 expressed genes (DEGs) were identified using MAST. For identifying DEGs in MOGHE-WM vs
200 Control-WM, MAST was applied using the number of counts, mitochondrial percentage and
201 dataset as latent variables. For MOGHE-WM vs MOGHE-GM, we used the number of counts,
202 mitochondrial percentage, donor, and brain region as latent variables. Genes with an adjusted P
203 value of less than 0.05 and fold-change threshold of 2 were considered significant. Mitochondrial
204 genes and those located in sex chromosomes were removed from the analysis.

205 Gene Ontology (GO) and Disease Ontology (DO) enrichment was assessed with clusterProfiler[26]
206 v.4.2.2 and DOSE[27] v.3.20.1 R Packages using the enrichGO and enrichDO functions,
207 respectively.

208

209 **Integration with snRNA-seq from heterotypical white matter**

210 We incorporated control white matter snRNA-seq data from two autopsy donors with no
211 neurological history, originally sequenced in Elkjaer et al.[28]. Raw sequencing data were obtained
212 from GEO (accession number GSE231585) and processed with cellranger count v.7.0.1 using the
213 human genome reference GChR38. Nuclei were filtered using Seurat based on the following
214 criteria: 1000 < UMI counts < 25,000, 200 < nFeatures < 4000, and mitochondrial read percentage
215 (percent.mt) < 1. This dataset was integrated into our MOGHE dataset using the procedure outlined
216 in the "snRNA-seq analysis" section.

217

218 **snATAC-seq analysis**

219 Samples were preprocessed in Signac to filter out low-quality nuclei based on these criteria: 1,000
220 < ATAC fragments < 50,000, ATAC fragments in peaks > 400, nucleosome signal < 2, and TSS
221 enrichment > 2. Peaks within each sample were identified using MACS2[29]. Standard processing
222 and normalization were applied with the following functions: RunTFIDF, FindTopFeatures,
223 RunSVD, and FindClusters. The peak counts matrix for each sample was normalized using
224 RunTFIDF to adjust for differences in sequencing depth. Dimensionality reduction was performed
225 with latent semantic indexing (LSI) via the RunSVD function, with UMAP projection utilizing
226 LSI components 2-30, as the first LSI component was determined to represent technical variation.
227 Prior to sample integration, a common peak set across samples was created using MACS2, and
228 peaks in each sample were quantified with the FeatureMatrix function. Peaks located in non-
229 standard chromosomes and blacklisted regions were excluded from subsequent analyses. We then
230 re-computed RunTFIDF, FindTopFeatures, and RunSVD on the merged dataset, and applied
231 Harmony[25] to integrate low-dimensional cell embeddings (LSI components 2-30) across samples,
232 using parameters similar to those defined for RNA analysis. Chromatin accessibility at the gene
233 level was computed with the GeneActivity function from the Signac package, which sums the
234 fragments intersecting the gene body and promoter regions of each gene.

235

236 **Multimodal data integration**

237 To integrate gene expression and chromatin accessibility data, we utilized Weighted Nearest
238 Neighbor (WNN) analysis, an unsupervised method that evaluates the contribution of each
239 modality and constructs a combined graph representing both RNA and ATAC data. WNN analysis
240 was conducted using the FindMultiModalNeighbors function with 20 neighbors (k.nn), based on
241 UMAP reductions obtained after Harmony[25] integration. The WNN graph was then employed to
242 generate a joint UMAP visualization. Clustering based on the WNN graph was performed using
243 the FindClusters function with the smart local moving (SLM) algorithm. Clusters with fewer than
244 50 nuclei, a high percentage of doublets, or significant enrichment in mitochondrial markers were
245 excluded from further analysis.

246

247 **Cluster annotation**

248 Cell type annotation was performed using Azimuth[20], with the Allen Human Motor Cortex Atlas[30]
249 as the reference dataset. Azimuth performed label transfer by assigning each cell in the MOGHE
250 dataset to the most similar cell type in the reference based on gene expression similarity. The cell
251 type assigned to each cluster was determined by identifying the cell type with the highest mean
252 frequency within that cluster. The Azimuth annotation was manually validated using marker genes,
253 as depicted in Figure 1E. Cluster markers were identified using the FindAllMarkers function on
254 normalized data. Differentially expressed markers were determined with the Wilcox rank-sum test,
255 applying a log2 fold-change threshold of 0.25 and an adjusted P value of less than 0.05.

256

257 **Differential cellular abundance**

258 We utilized the permutation_test function from the scProportionTest R package[31] v. 0.0.0.9 to
259 assess changes in cell type composition between tissue types (MOGHE-WM *vs* Control-WM and
260 MOGHE-WM *vs* MOGHE-GM). The significance test used 10,000 permutations, and cell type
261 changes with a false discovery rate (FDR) less than 0.05 and a fold difference greater than 2 were
262 considered statistically significant.

263

264 **Gene signature analysis**

265 Gene signature scores were computed using the AddModuleScore function from the Seurat
266 package applied to log-normalized RNA data. Signatures for oligodendrocyte subtypes were
267 obtained from Jakel et al., 2019[32] and Sadick et al., 2022[33].

268

269 **Motif enrichment analysis**

270 We utilized chromVAR v3.3.2[34] and Signac to perform motif analysis. Nuclei-level motif activity
271 was computed using the RunChromVAR function in Signac with the set of background peaks
272 matched appropriately. Differential activity of motifs was determined using the FindMarkers
273 function, applying a likelihood ratio (LR) test with ATAC fragment counts as a latent variable, a
274 minimum percentage of cells expressing the motif (min.pct) set to 0.05, and an adjusted p-value
275 cutoff of < 0.05. To ensure biologically-relevant results, motifs were further filtered based on the
276 expression of their associated regulators, which were required to be expressed in at least 10% of
277 the cells within a given cluster/cell type.

278

279 **RESULTS**

280

281 **Multimodal single-nucleus profiling of MOGHE lesions**

282 We obtained frozen brain specimens from two donors aged 22 and 57 years with a confirmed
283 pathological diagnosis of MOGHE after epilepsy surgery (Table 1). After isolating nuclei from the
284 MOGHE gray matter (MOGHE-GM) and white matter regions (MOGHE-WM), we performed
285 paired snRNA-seq and snATAC-seq on each nucleus using the 10X Genomics Multiome assay to
286 measure chromatin accessibility and gene expression at nuclei resolution (Figure 1A). The
287 sequenced tissues fulfilled the histopathological criteria for MOGHE, presenting with increase of
288 the oligodendrocyte density in the WM and excess of heterotopic neurons (Figure 1B, Methods).
289 Deep exome sequencing of the donors using the MOGHE-WM has not detected somatic variants
290 in *SLC35A2* or in genes previously reported in epileptogenic brain lesions[35].

291 After filtering to remove low-quality nuclei and doublets (Methods), we obtained 31,659 nuclei
292 derived from MOGHE samples in the multiome snATAC-seq + snRNA-seq dataset (Figure 1C).
293 Next, we integrated our MOGHE dataset with snRNA-seq performed on white matter from two
294 neurotypical individuals (Control-WM) obtained from Elkjaer et al.[28], resulting in a total of 39,743
295 nuclei (Figure 1D, Table 1). Clustering analysis identified 33 distinct clusters after integration of
296 MOGHE and control nuclei (Figure S1A). Cell type annotation was carried out using Azimuth[20],
297 and subsequently validated through manual inspection of canonical marker genes (Figure 1E). The
298 Azimuth annotation enabled confident labeling of most of these clusters (Figure S1B), and most
299 clusters were shared across multiple samples (Figure S1C). After annotation, the majority of nuclei
300 was classified as oligodendrocytes (OLs), followed by excitatory neurons (ENs, further divided
301 into upper- and deep-layer subsets), astrocytes, inhibitory neurons (INs, subclassified into Vip,
302 Sst, Pvalb, Lamp5 and Scng subtypes), oligodendrocyte precursor cells (OPCs), microglia and
303 endothelial/vascular and leptomeningeal cells (Endo/VLMCs). As expected, samples derived from
304 white matter (MOGHE-WM and Controls-WM) were predominantly composed of OLs, while
305 neurons were not detected in Controls-WM and primarily located within MOGHE-GM (Figure
306 S1D).

307 We also examined gene accessibility profiles computed based on the chromatin accessibility
308 (snATAC-seq) for the same set of markers (Figure 1F). Gene accessibility displayed a strong
309 correlation with gene expression across cell type markers, with the exception of some endothelial
310 markers, denoting the agreement between data modalities. Thus, our single-nucleus profiling
311 identified the various cortical cell types in GM and subcortical WM regions, revealing the cellular
312 diversity of MOGHE lesions.
313

314 **Cellular and gene expression changes in MOGHE white matter**

315 To investigate disease-associated changes in the WM, we analyzed cell type proportions between
316 MOGHE-WM and Controls-WM as well as MOGHE-WM and MOGHE-GM. Compared to
317 Controls-WM, MOGHE-WM tissue had a higher proportion of ENs, INs and microglia, denoting
318 the expansion of heterotopic neurons and OLs in the WM of the abnormal tissue (Figure 2A). Also,
319 MOGHE-WM had an increased proportion of OLs and a decreased abundance of astrocytes, ENs,

320 and INs compared to MOGHE-GM, reflecting the expected differences in cell composition
321 between cortical and subcortical regions (Figure 2B).

322 We also analyzed differentially expressed genes (DEGs) between tissue conditions and within each
323 cell type. Neurons were not considered in this analysis since they were not identified in Controls-
324 WM. In MOGHE-WM compared to Controls-WM, there were significant gene expression changes
325 in most glial populations, including astrocytes, microglia, and OLs (Figure 2C). When compared
326 to MOGHE-GM, OLs in MOGHE-WM were the most affected population with a large number of
327 downregulated genes (Figure 2C). Collectively, these analyses indicate that the cellular
328 composition of subcortical MOGHE-WM was characterized by an increased OL population and
329 the presence of heterotopic neurons, which recapitulated key hallmarks of the disease. Meanwhile,
330 gene expression changes revealed that nearly all glial cell types in subcortical MOGHE-WM were
331 affected by the condition.

332

333 **Pathological oligodendrocyte states in MOGHE**

334 To identify MOGHE-associated OLs, we performed a subcluster analysis in the OLs compartment
335 and assessed whether these subpopulations were enriched in disease tissue. Notably, subclusters 4
336 and 7 (OL_4 and OL_7), which clustered separately from the major OLs population, were
337 predominantly observed in MOGHE (Figure 3A). Differential cellular abundance analysis
338 confirmed that OL_4 and OL_7 were significantly expanded in MOGHE-WM compared to
339 Controls-WM (Figure 3B, top). OL_4 was also present in MOGHE-GM, while OL_7 was mostly
340 derived from MOGHE-WM (Figure 3B, bottom).

341 MOGHE-associated OLs expressed marker genes related to immune genes (*C3*, *CD74*, and *CD83*),
342 myelination disruption (*APOE4*), and OL differentiation (*FGFR3*) (Figure 3C, Supplemental
343 Table 1). In addition, *OLIG2*, a disease marker used to identify OLs clusters in MOGHE tissue in
344 neuropathological routine, was exclusively expressed in OL_7 subpopulation (Figure 3D), even
345 though the chromatin in *OLIG2* promoter region was accessible across all subclusters (Figure 3E).

346 Further, we performed a gene ontology (GO) analysis on OL_4 and OL_7 markers to identify
347 molecular functions associated with these MOGHE-specific OLs. Enriched GO terms included

348 synaptic organization and assembly, regulation of trans-synaptic signaling, regulation of
349 membrane potential, neurotransmitter support, and axonogenesis (Figure 3F, Supplemental Table
350 1). Functional enrichment using the Disease Ontology showed that OL_4 and OL7 subpopulations
351 expressed marker genes related to brain diseases, including epileptic and neuropsychiatric
352 disorders (Figure 3G, Supplemental Table 1).

353 Previous single-nucleus studies identified OLs subsets associated with neurological conditions,
354 including the immune subtype (ImOLGs) and the integrated 2 subtype (Int2), which were
355 identified in multiple sclerosis and Alzheimer's disease, respectively[32,33]. Based on this
356 observation, we conducted a signature analysis to correlate these pathological OLs signatures with
357 OLs identified in MOGHE lesions. Notably, OL_4 and OL_7 also exhibited the highest scores for
358 ImOLGs (Figures 3H,J) and Int2 (Figures 3I,J) subsets, indicating that MOGHE-specific OL_4
359 and OL_7 share a common signature with OLs identified in other neurological diseases. In
360 addition, OL_4 and OL_7 expressed genes linked to epilepsy[36] such as *SLC1A2*, *KNCQ2*, *KCNQ3*
361 (Figure 3J). OL_7 expressed markers previously reported in diseased-associated oligodendrocytes
362 in mouse models of MS and AD[37] including *ATF4*, *EGR1*, and *STAT1* (Figure 3J). Altogether,
363 these analyses characterized the diverse OLs subpopulations in MOGHE and enabled us to identify
364 perturbed OL cell states in the disease.

365 **Characterization of heterotopic neurons in MOGHE**

366 We identified a subpopulation of heterotopic neurons (HNs) derived specifically from MOGHE-
367 WM samples (n = 246; Figures 4A,B), which were absent in both Control-WM and MOGHE-GM
368 samples. Most of the HNs were annotated as ENs from layers 5/6 (n = 129; 52%), and the
369 remaining nuclei consisted mainly of INs subtypes Lamp5, Sst and PvalB. The HNs subpopulation
370 expressed a very distinct set of markers such as cell-adhesion genes (*CDH5* and *CDH3*),
371 transmembrane proteins (*TMEM275* and *TMEM233*), ion channels (*KCNK2*, *CACNA1B*), and the
372 neurotransmitter *NPY* (Figure 4C, Supplemental Table 2).

373 To identify molecular functions in this abnormal neuronal subpopulation, we analyzed DEGs
374 between HNs and total MOGHE-GM neurons, identifying 255 upregulated and 233 downregulated
375 genes (Figure 4D, Supplemental Table 2). Enriched GO terms associated with downregulated
376 genes included negative regulation of neuron projection, cytoplasmic translation, and cell junction
377 assembly (Figure 4E). On the other hand, upregulated molecular functions included positive

378 regulation of cell projection, the Fc receptor-mediated stimulatory signaling pathway, and the Wnt
379 signaling pathway (Figure 4F), which has been shown to regulate neuronal migration[38,39].
380 In line with the DEG analysis, a motif enrichment analysis based on the snATAC-seq data
381 identified that FOX transcriptional regulators were among the top activated motifs in HNs (Figure
382 4F). The FOX family has been shown to regulate the Wnt signaling pathway[40,41], which could
383 explain the distinctive Wnt activation in HNs. Therefore, these results indicate that HNs in
384 MOGHE consisted of ENs from deep cortical layers with increased expression of pathways
385 associated with neuronal migration.
386
387 **DISCUSSION**
388
389 In this study, we applied multiomic single-nucleus profiling to systematically map the cellular
390 landscapes of cortical and subcortical regions affected in MOGHE, a type of focal MCD associated
391 with epilepsy. Overall, we identified significant changes in cell type composition in MOGHE-WM
392 and uncovered distinct pathological cell populations in the disease.

393 We identified MOGHE-associated OLs subpopulations present in WM tissue, overexpressing
394 genes involved in immune response and myelination disruption. Among these genes, APOE4 has
395 been shown to impair oligodendrocyte differentiation[42] and to downregulate myelination via
396 cholesterol dysregulation[43]. Cholesterol biosynthesis is essential to myelin biogenesis[44,45]. Thus,
397 these dysregulated OLs states might contribute to WM matter degradation and the formation of
398 irregular zones of hypomyelination typically observed in MOGHE-WM. Noteworthy, we
399 demonstrated that these MOGHE-associated OLs share a common transcriptional signature with
400 OLs identified in multiple sclerosis and Alzheimer's disease, which are neurological conditions
401 presenting with WM degradation and myelin loss[46,47].

402 Other glial cell types may also contribute to the disease, as evidenced by microglia increase in
403 MOGHE-WM. Microglia have been linked to epilepsy and their effects can be pro- or anti-
404 inflammatory depending on the disease stage[48,49]. Moreover, resident microglia are required for
405 maintaining myelin health and their dysregulation may be related to the emergence of pathological
406 OL states[50].
407

We also identified the appearance of misplaced neurons in the WM (HNs), another key pathological hallmark of MOGHE. We found that MOGHE HNs displayed increased expression of neuronal projection genes and Wnt signaling pathway. Interestingly, Wnt signaling has been linked to axonal/dendrite remodeling and neuronal plasticity[51,52], and regulates neuronal migration by modulating the microtubule cytoskeleton[38,39], providing a potential mechanism for the abnormal localization of HNs. Furthermore, alterations in Wnt signaling induce neuronal hyperexcitability[53,54], a defining feature of epilepsy. Thus, the HN population may contribute to the development of epileptogenic circuits in MOGHE.

In summary, this high-resolution mapping of MOGHE lesions in clinical samples has generated the first cellular atlas of human tissue affected by the disease, providing a more detailed view of the dysregulated cell populations involved in this pathology.

Author contributions

D.F.T.V. and F.R. designed the study. M.L. performed single-nucleus assays. I.C.G. and D.F.T.V. performed data analyses. C.L.Y., E.G., H.T., M.K.M.A., F.C., and I.L-C contributed to clinical sample collection. L.K., I.B. and F.R. performed histopathological analyses. I.C.G. and D.F.T.V. wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

Funding Sources

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) [grant numbers 2019/07382-2, 2022/01530-2, 2019/08259-0, 2013/07559-3]; the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation [grant number DAF2021-237598]; Conselho Nacional de Pesquisa (CNPq) [grant number 311923/2019-4].

Declaration of interests

The authors declare no competing interests.

437 **REFERENCES**
438

439 1. Severino M, Geraldo AF, Utz N, Tortora D, Pogledic I, Klonowski W, et al. Definitions and
440 classification of malformations of cortical development: practical guidelines. Brain. 2020;
441 143(10):2874–94.

442 2. Schurr J, Coras R, Rössler K, Pieper T, Kudernatsch M, Holthausen H, et al. Mild
443 Malformation of Cortical Development with Oligodendroglial Hyperplasia in Frontal Lobe
444 Epilepsy: A New Clinico-Pathological Entity. Brain Pathology. 2017; 27(1):26–35.

445 3. Blümcke I. Neuropathology and epilepsy surgery: 2022 update. Vol. 3, Free Neuropathology.
446 University of Muenster; 2022.

447 4. Najm I, Lal D, Alonso Vanegas M, Cendes F, Lopes-Cendes I, Palmini A, et al. The ILAE
448 consensus classification of focal cortical dysplasia: An update proposed by an ad hoc task
449 force of the ILAE diagnostic methods commission. Epilepsia. 2022; 63(8):1899–919.

450 5. Mendes Coelho VC, Morita-Sherman M, Yasuda CL, Alvim MMK, Amorim BJ, Tedeschi H,
451 et al. Magnetic resonance imaging findings and clinical characteristics in mild
452 malformation of cortical development with oligodendroglial hyperplasia and epilepsy in a
453 predominantly adult cohort. Epilepsia. 2021; 62(6):1429–41.

454 6. Gaballa A, Woermann FG, Cloppenborg T, Kalbhenn T, Blümcke I, Bien CG, et al. Clinical
455 characteristics and postoperative seizure outcome in patients with mild malformation of
456 cortical development and oligodendroglial hyperplasia. Epilepsia. 2021; 62(12):2920–31.

457 7. Bonduelle T, Hartlieb T, Baldassari S, Sim NS, Kim SH, Kang H-C, et al. Frequent SLC35A2
458 brain mosaicism in mild malformation of cortical development with oligodendroglial
459 hyperplasia in epilepsy (MOGHE). Acta Neuropathologica Communications. 2021; 9(1):3.

460 8. López-Rivera JA, Leu C, Macnee M, Khoury J, Hoffmann L, Coras R, et al. The genomic
461 landscape across 474 surgically accessible epileptogenic human brain lesions. Brain. 2022;
462 146(4):1342.

463 9. Elziny S, Sran S, Yoon H, Corrigan RR, Page J, Ringland A, et al. Loss of *Slc35a2* alters
464 development of the mouse cerebral cortex. Neuroscience Letters. 2024; 836:137881.

465 10. Spyrou J, Aung KP, Vanyai H, Leventer RJ, Maljevic S, Lockhart PJ, et al. *Slc35a2* mosaic
466 knockout impacts cortical development, dendritic arborisation, and neuronal firing.
467 Neurobiology of Disease. 2024; 201:106657.

468 11. Yoon H, Ringland A, Anderson JJ, Sran S, Elziny S, Huynh C, et al. Mouse models of
469 Slc35a2 brain mosaicism reveal mechanisms of mild malformations of cortical
470 development with oligodendroglial hyperplasia in epilepsy. Epilepsia. 2024; .

471 12. Liddelow SA, Guttenplan KA, Clarke LE, Bennett FC, Bohlen CJ, Schirmer L, et al.
472 Neurotoxic reactive astrocytes are induced by activated microglia. Nature. 2017;
473 541(7638):481–7.

474 13. Olah M, Menon V, Habib N, Taga MF, Ma Y, Yung CJ, et al. Single cell RNA sequencing
475 of human microglia uncovers a subset associated with Alzheimer's disease. Nat Commun.
476 2020; 11(1):6129.

477 14. Adams L, Song MK, Yuen S, Tanaka Y, Kim Y-S. A single-nuclei paired multiomic
478 analysis of the human midbrain reveals age- and Parkinson's disease–associated glial
479 changes. Nature aging. 2024; 4(3):364.

480 15. Galvão IC, Lemoine M, Messias LA, Araújo PAORA, Geraldis JC, Yasuda CL, et al.
481 Multimodal single-cell profiling reveals neuronal vulnerability and pathological cell states
482 in focal cortical dysplasia. iScience 2024; 27(12).

483 16. Walter TJ, Suter RK, Ayad NG. An overview of human single-cell RNA sequencing studies
484 in neurobiological disease. Neurobiology of Disease. 2023; 184:106201.

485 17. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform.
486 Bioinformatics. 2009; 25(14):1754–60.

487 18. DePristo MA, Banks E, Poplin RE, Garimella KV, Maguire JR, Hartl C, et al. A framework
488 for variation discovery and genotyping using next-generation DNA sequencing data. Nat
489 Genet. 2011; 43(5):491–8.

490 19. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for
491 annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly
492 (Austin). 2012; 6(2):80–92.

493 20. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated
494 analysis of multimodal single-cell data. Cell. 2021; 184(13):3573-3587.e29.

495 21. Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis
496 with Signac. Nat Methods. 2021; 18(11):1333–41.

497 22. Yang S, Corbett SE, Koga Y, Wang Z, Johnson WE, Yajima M, et al. Decontamination of
498 ambient RNA in single-cell RNA-seq with DecontX. Genome Biology. 2020; 21(1):57.

499 23. celda [Internet]. Bioconductor. [cited 2024]. Available from:
500 http://bioconductor.org/packages/celda/

501 24. Germain P-L, Lun A, Garcia Meixide C, Macnair W, Robinson MD. Doublet identification
502 in single-cell sequencing data using scDblFinder. F1000Res. 2021; 10:979.

503 25. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and
504 accurate integration of single-cell data with Harmony. Nat Methods. 2019; 16(12):1289–96.

505 26. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment
506 tool for interpreting omics data. Innovation [Internet]. 2021 [cited 2024]; 2(3). Available
507 from: https://www.cell.com/the-innovation/abstract/S2666-6758(21)00066-7

508 27. Yu G, Wang L-G, Yan G-R, He Q-Y. DOSE: an R/Bioconductor package for disease
509 ontology semantic and enrichment analysis. Bioinformatics. 2015; 31(4):608–9.

510 28. Elkjaer ML, Hartebrodt A, Oubounyt M, Weber A, Vitved L, Reynolds R, et al. Single-Cell
511 Multi-Omics Map of Cell Type–Specific Mechanistic Drivers of Multiple Sclerosis
512 Lesions. Neurology Neuroimmunology & Neuroinflammation. 2024; 11(3):e200213.

513 29. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based
514 Analysis of ChIP-Seq (MACS). Genome Biology. 2008; 9(9):R137.

515 30. Bakken TE, Jorstad NL, Hu Q, Lake BB, Tian W, Kalmbach BE, et al. Comparative
516 cellular analysis of motor cortex in human, marmoset and mouse. Nature. 2021;
517 598(7879):111–9.

518 31. Miller SA, Policastro RA, Sriramkumar S, Lai T, Huntington TD, Ladaika CA, et al. LSD1
519 and Aberrant DNA Methylation Mediate Persistence of Enteroendocrine Progenitors That
520 Support BRAF-Mutant Colorectal Cancer. Cancer Res. 2021; 81(14):3791–805.

521 32. Jäkel S, Agirre E, Mendanha Falcão A, van Bruggen D, Lee KW, Knuesel I, et al. Altered
522 human oligodendrocyte heterogeneity in multiple sclerosis. Nature. 2019; 566(7745):543–
523 7.

524 33. Sadick JS, O'Dea MR, Hasel P, Dykstra T, Faustin A, Liddelow SA. Astrocytes and
525 oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease.
526 Neuron. 2022; 110(11):1788-1805.e10.

527 34. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-
528 associated accessibility from single-cell epigenomic data. Nature Methods. 2017;
529 14(10):975–8.

530 35. López-Rivera JA, Leu C, Macnee M, Khoury J, Hoffmann L, Coras R, et al. The genomic
531 landscape across 474 surgically accessible epileptogenic human brain lesions. Brain. 2023;
532 146(4):1342–56.

533 36. Macnee M, Pérez-Palma E, López-Rivera JA, Ivaniuk A, May P, Møller RS, et al. Data-
534 driven historical characterization of epilepsy-associated genes. European Journal of
535 Paediatric Neurology. 2023; 42:82–7.

536 37. Pandey S, Shen K, Lee S-H, Shen Y-AA, Wang Y, Otero-García M, et al. Disease-
537 associated oligodendrocyte responses across neurodegenerative diseases. Cell Reports.
538 2022; 40(8):111189.

539 38. Hayat R, Manzoor M, Hussain A. Wnt signaling pathway: A comprehensive review. Cell
540 Biology International. 2022; 46(6):863–77.

541 39. Salinas PC. Modulation of the microtubule cytoskeleton: a role for a divergent canonical
542 Wnt pathway. Trends in Cell Biology. 2007; 17(7):333–42.

543 40. den Hoed J, Devaraju K, Fisher SE. Molecular networks of the FOXP2 transcription factor
544 in the brain. EMBO reports. 2021; 22(8):e52803.

545 41. Moparthi L, Koch S. FOX transcription factors are common regulators of Wnt/β-catenin–
546 dependent gene transcription. Journal of Biological Chemistry [Internet]. 2023 [cited 2024];
547 299(5). Available from: https://www.jbc.org/article/S0021-9258(23)00309-5/abstract

548 42. Mok KK-S, Yeung SH-S, Cheng GW-Y, Ma IW-T, Lee RH-S, Herrup K, et al.
549 Apolipoprotein E ε4 disrupts oligodendrocyte differentiation by interfering with astrocyte-
550 derived lipid transport. J Neurochem. 2023; 165(1):55–75.

551 43. Blanchard JW, Akay LA, Davila-Velderrain J, von Maydell D, Mathys H, Davidson SM, et
552 al. APOE4 impairs myelination via cholesterol dysregulation in oligodendrocytes. Nature.
553 2022; 611(7937):769–79.

554 44. Berghoff SA, Gerndt N, Winchenbach J, Stumpf SK, Hosang L, Odoardi F, et al. Dietary
555 cholesterol promotes repair of demyelinated lesions in the adult brain. Nat Commun. 2017;
556 8(1):14241.

557 45. Saher G, Brügger B, Lappe-Siefke C, Möbius W, Tozawa R, Wehr MC, et al. High
558 cholesterol level is essential for myelin membrane growth. Nat Neurosci. 2005; 8(4):468–
559 75.

560 46. Nasrabady SE, Rizvi B, Goldman JE, Brickman AM. White matter changes in Alzheimer's
561 disease: a focus on myelin and oligodendrocytes. Acta Neuropathologica Communications.
562 2018; 6(1):22.

563 47. Ontaneda D, Hyland M, Cohen JA. Multiple Sclerosis: New Insights in Pathogenesis and
564 Novel Therapeutics. Annual Review of Medicine. 2012; 63(Volume 63, 2012):389–404.

565 48. Yu C, Deng X, Xu D. Microglia in epilepsy. Neurobiology of Disease. 2023; 185:106249.

566 49. Olah M, Menon V, Habib N, Taga MF, Ma Y, Yung CJ, et al. Single cell RNA sequencing
567 of human microglia uncovers a subset associated with Alzheimer's disease. Nat Commun.
568 2020; 11(1):6129.

569 50. McNamara NB, Munro DAD, Bestard-Cuche N, Uyeda A, Bogie JFJ, Hoffmann A, et al.
570 Microglia regulate central nervous system myelin growth and integrity. Nature. 2023;
571 613(7942):120–9.

572 51. He C-W, Liao C-P, Pan C-L. Wnt signalling in the development of axon, dendrites and
573 synapses. Open Biology. 2018; 8(10):180116.

574 52. McLeod F, Salinas PC. Wnt proteins as modulators of synaptic plasticity. Current Opinion
575 in Neurobiology. 2018; 53:90–5.

576 53. Jaworski T. Control of neuronal excitability by GSK-3beta: Epilepsy and beyond.
577 Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2020; 1867(9):118745.

578 54. Priya, Yadav N, Anand S, Banerjee J, Tripathi M, Chandra PS, et al. The multifaceted role
579 of Wnt canonical signalling in neurogenesis, neuroinflammation, and hyperexcitability in
580 mesial temporal lobe epilepsy. Neuropharmacology. 2024; 251:109942.

581

582

583 **TABLES WITH TITLES AND LEGENDS**

584

585 **Table 1** - Description of the clinical samples.

Condition	Specimen identifier	Donor	Dissection	Sex	Age at surgery (y)	Age at onset (y)	Lobe	Side of resection	Tissue source	Sequenced nuclei
MOGHE (This study)	G23_GM	G23	GM	F	22	11	Frontal	Right	Epilepsy surgery	4,603
	G23_WM	G23	WM	F	22	11	Frontal	Right	Epilepsy surgery	10,804
	G39_GM	G39	GM	M	57	36	Temporal	Left	Epilepsy surgery	1,593
	G23_WM	G39	WM	M	57	36	Temporal	Left	Epilepsy surgery	14,659
Controls (Elkjaer *et al.*, 2024)	C3	C3	WM	M	68	NA	NA	NA	Post-mortem	3,810
	C4	C4	WM	F	50	NA	NA	NA	Post-mortem	4,274

586

587 Abbreviations: F, female; M, male; NA, not available.

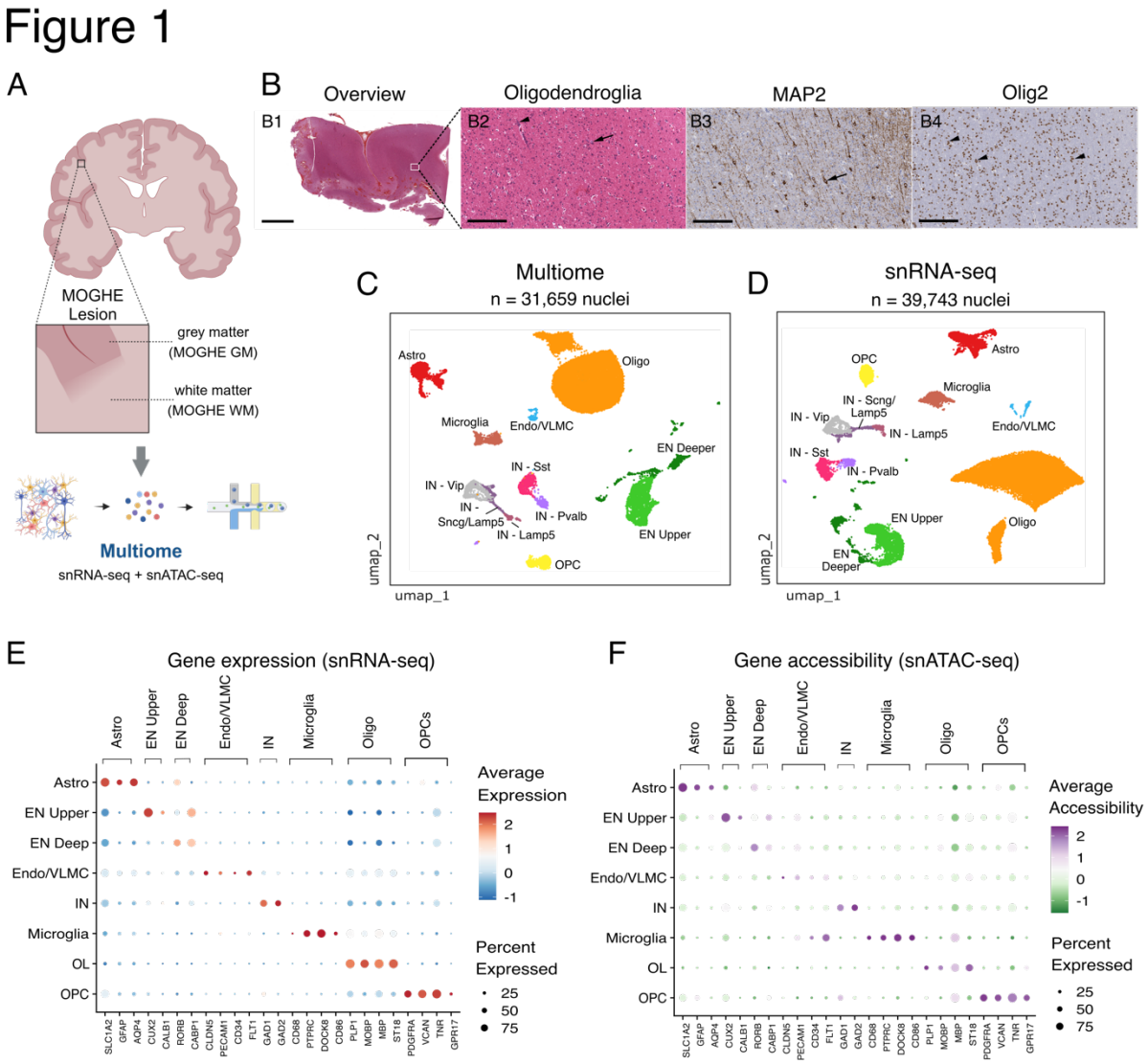
588

589

590

591

592 **FIGURES**



593

594 **Figure 1. Multimodal single-nucleus sequencing of MOGHE tissue.**

595 **(A)** Schematic of MOGHE samples and study design. Nuclei from the MOGHE grey matter
596 (MOGHE-GM) and white matter regions (MOGHE-WM) isolated and profiled by snRNA-seq
597 and snATAC-seq using the Multiome assay (10X Genomics). Created with Biorender.

598 **(B)** Representative histopathological findings of a brain surgical sample with MOGHE sequenced
599 in the study. **B1.** Hematoxylin and eosin (H&E) stained section showing an overview of the
600 resected tissue. The white rectangle indicates a subcortical WM region near the GM-WM junction

601 presented in B2-B4. **B2.** Clusters of oligodendroglial cells either perivascular (arrow head) or
602 around heterotopic neurons (arrow). **B3.** Immunohistochemistry (IHC) for MAP2 (neurons)
603 depicts scattered heterotopic neurons (arrow). **B4.** IHC for Olig2 (oligodendrocytes) showing an
604 irregularly distributed increase of the oligodendroglial cells, either isolated or in clusters (arrow
605 heads). Scale bars: 4 mm (Figure B1) and 200 µm (Figures B2-B4).
606 **(C)** Joint uniform manifold approximation and projection (UMAP) visualization of snATAC-seq
607 and snRNA-seq from nuclei sequenced in MOGHE samples, colored by annotated cell type.
608 **(D)** UMAP visualization of snRNA-seq from MOGHE samples integrated with neurotypical WM
609 controls obtained from Elkjaer et al.[28], and colored by cell type.
610 **(E)** Dot plot displaying gene expression levels for canonical markers of major cortical brain cell
611 types. The size of each dot represents the proportion of cells expressing the marker, while the color
612 indicates the average expression levels.
613 **(F)** Dot plot displaying gene accessibility levels computed on the basis of chromatin accessibility
614 for the same marker genes shown in (E). The size of each dot represents the proportion of cells
615 expressing the marker, while the color indicates the average gene accessibility.
616 Astro, astrocytes. EN, excitatory neurons. Endo/VLMC, endothelial/vascular and leptomeningeal
617 cells. IN, inhibitory neurons. OL, oligodendrocytes. OPC, oligodendrocyte precursor cells.
618

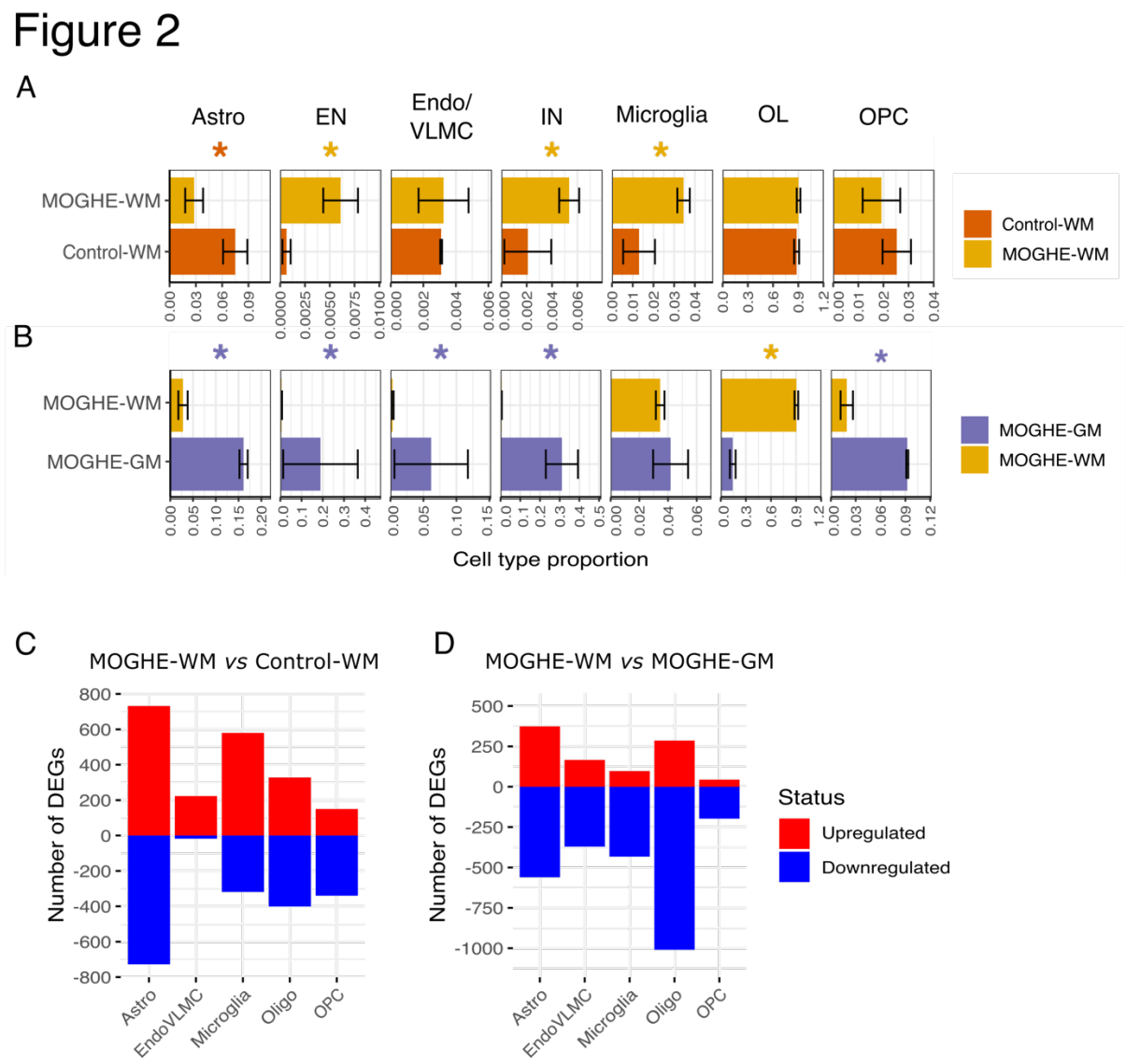


Figure 2. Cell type abundance and gene expression changes in MOGHE.

(A) Barplots denoting cell types abundance in MOGHE-WM and Control-WM samples. The *x*-axis indicates the cell type proportion, and the *y*-axis indicates the tissue type. Error bars indicate the standard error of the mean. Differential cell type abundance was performed using the scProportionTest tool (Methods). Asterisks indicate significant changes considering an FDR < 0.05 and fold-change > 1.5. The asterisk color indicates the tissue type with the higher abundance. (B) Barplots denoting cell types abundance in MOGHE-WM and MOGHE-GM samples. Legends are defined as in (A).

628 (C) Differentially expressed genes (DEGs) between MOGHE-WM and Control-WM, in non-
629 neuronal populations. DEGs were obtained using MAST, with a fold-change > 2 and adjusted P
630 value < 0.05. (see Methods). The *x*-axis indicates the cell type, and the *y*-axis indicates the number
631 of DEGs.
632 (D) DEGs in MOGHE-WM vs MOGHE-GM. Legend is defined in (C).
633 Astro, astrocytes. EN, excitatory neurons. Endo/VLMC, endothelial/vascular and leptomeningeal
634 cells. IN, inhibitory neurons. OL, oligodendrocytes. OPC, oligodendrocyte precursor cells.
635
636

Figure 3

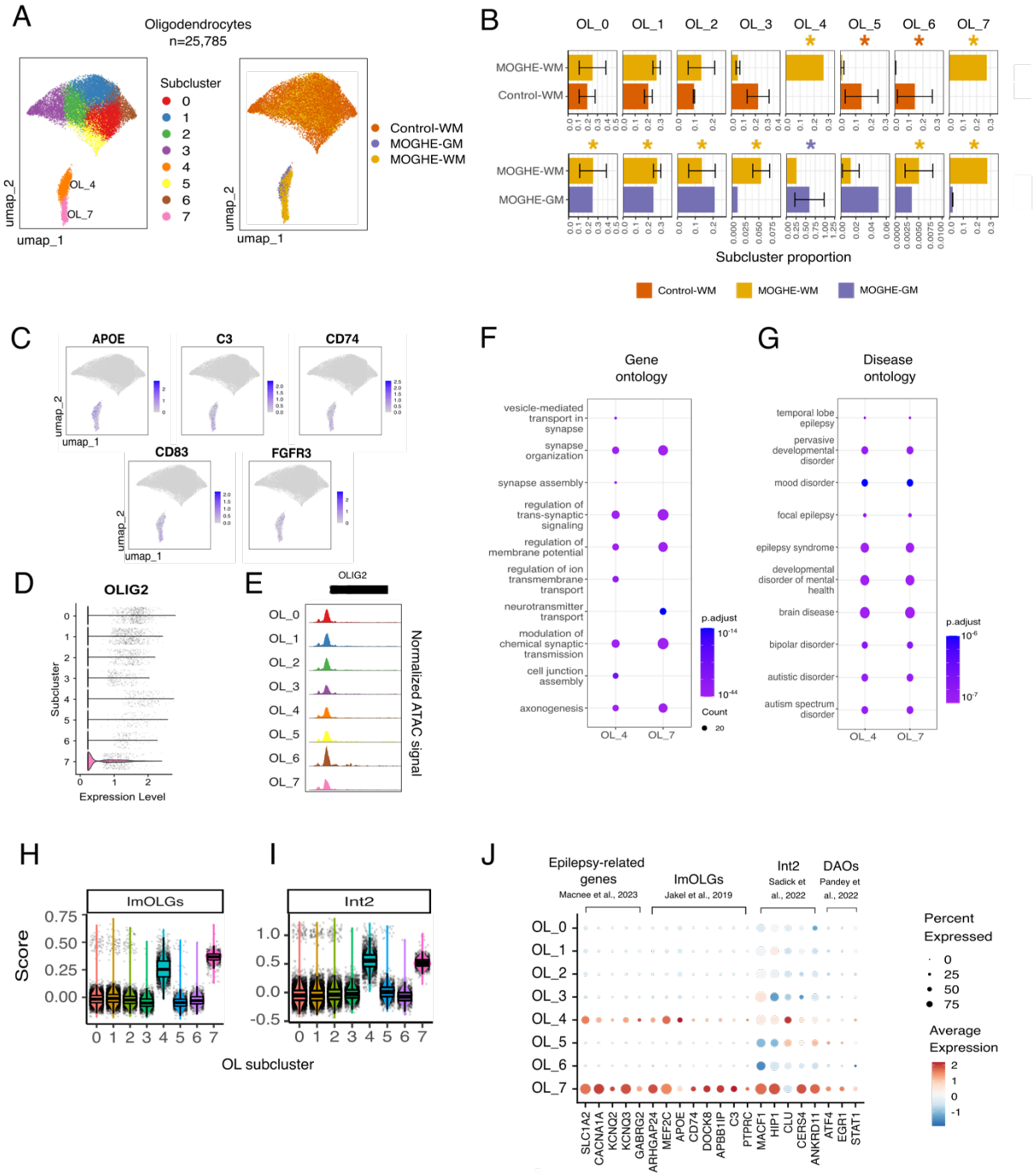


Figure 3. Identification of a MOGHE-specific oligodendrocyte population.

(A) UMAP representation of oligodendrocytes (OLs) from MOGHE-WM, MOGHE-GM and Control-WM, colored by subcluster (left) and tissue type (right).

641 **(B)** Bar plot showing the OL subcluster proportions in MOGHE-WM and Control-WM samples.
642 The *x*-axis indicates the subcluster proportion, and the *y*-axis indicates the tissue type. Differential
643 subcluster abundance was performed using the scProportionTest tool (Methods). Asterisks indicate
644 significant changes considering an FDR < 0.05 and fold-change > 2. The asterisk color indicates
645 the tissue type with the higher abundance. Error bars indicate standard error of the mean.
646 **(C)** Feature plots denoting *APOE*, *C3*, *CD74*, *CD83*, and *FGF3* gene expression in OLs. Cells are
647 colored according to normalized expression levels.
648 **(D)** Violin plot of *OLIG2* expression in OL subclusters. The *x*-axis indicates the subcluster and the
649 *y*-axis indicates the normalized expression level.
650 **(E)** Chromatin accessibility profile of *OLIG2* in OL subclusters from snATAC-seq. The
651 normalized ATAC signal is depicted in a region of +/− 1 kb around the gene start/end coordinates.
652 **(F)** Dot plot showing top enriched Gene Ontology (GO) terms in subclusters OL_4 and OL_7
653 computed using clusterProfiler[26]. The size of each dot represents the number of genes associated
654 with the GO term, while the color reflects the enrichment P-value.
655 **(G)** Dot plot showing top enriched Disease Ontology (DO) terms in subclusters OL_4 and OL_7
656 computed using DOSE[27] and clusterProfiler[26]. The size of each dot represents the number of genes
657 associated with the DO term, while the color reflects the enrichment P-value.
658 **(H)** Boxplots depicting activity scores of the immune oligodendroglia (ImOLG) signature from
659 Jakel et al[32] in OL subclusters. The *x*-axis indicates the subcluster, and the *y*-axis indicates the
660 activity score. The center line of the boxplot shows the median of the data; the box limits show the
661 upper and lower quartiles; the whiskers show 1.5 times interquartile ranges. Overlay dots represent
662 activity scores in individual nuclei.
663 **(I)** Boxplots depicting activity scores of the Int2 OL signature from Sadick et al[33] in subclusters.
664 Legends are defined in (H).
665 **(J)** Dot plot displaying expression levels of gene sets overexpressed in OL_4 and OL_7: epilepsy-
666 associated genes from Macnee et al.[36], genes from the ImOLG signature from Jakel et al.[32], genes
667 from the Int2 signature from Sadick et al.[33], and disease-associated oligodendrocytes (DAOs) from
668 Pandey et al.[37]. The size of each dot represents the proportion of cells expressing the gene in
669 subclusters, while the color indicates the normalized expression level.

Figure 4

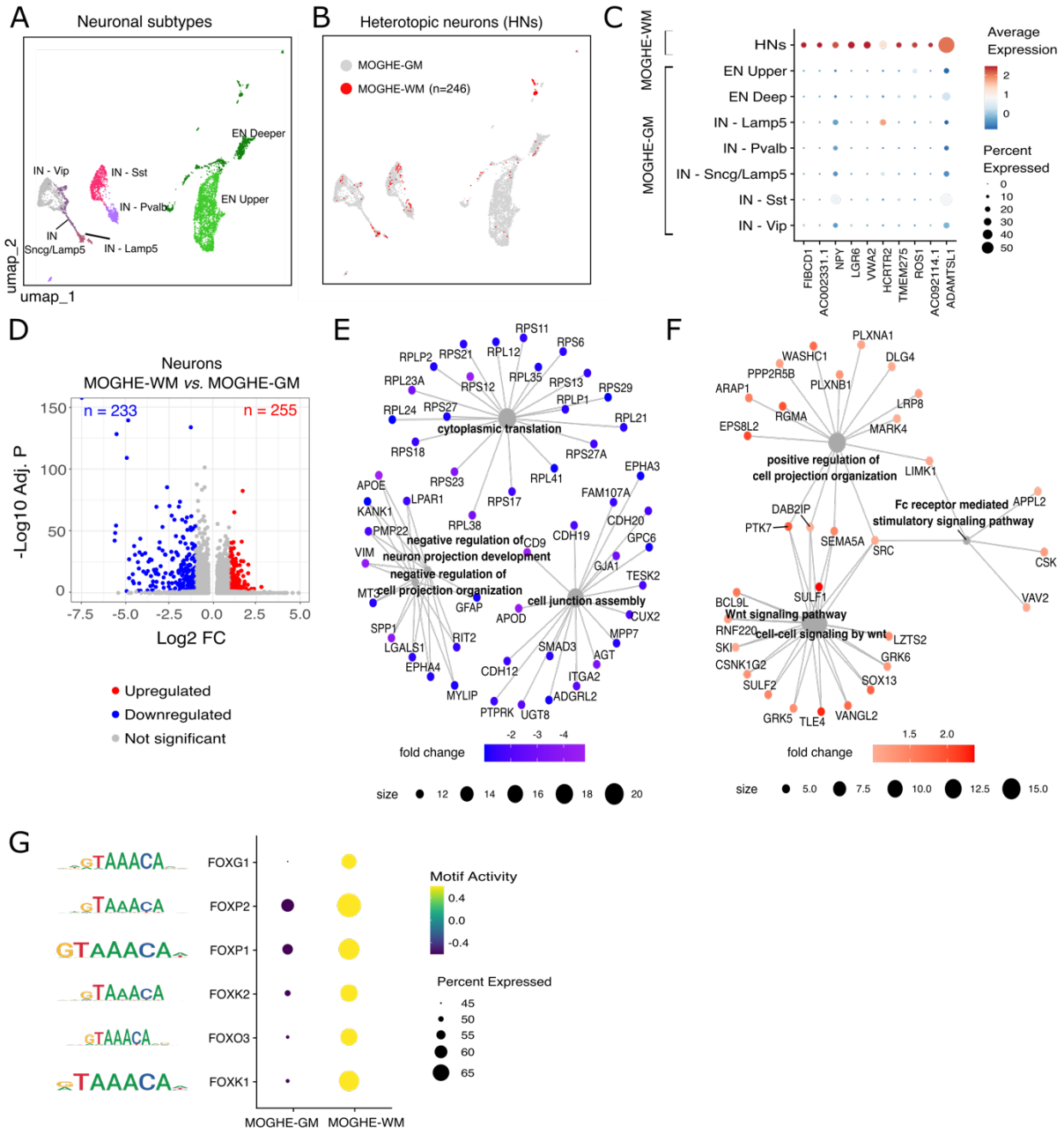


Figure 4. Characterization of heterotopic neurons in MOGHE.

(A) UMAP representation of neuronal nuclei from MOGHE-WM and MOGHE-GM, colored by neuronal subtype.

(B) UMAP representation of neuronal nuclei from (A), colored by tissue origin. Heterotopic neurons (HNs) are indicated in red.

676 **(C)** Dot plot displaying gene expression of HNs marker genes. The size of each dot represents the
677 proportion of cells expressing the marker, while the color indicates the average expression levels.
678 **(D)** Volcano plot displaying differentially expressed genes (DEGs) in MOGHE-WM *vs* MOGHE-
679 GM. DEGs were obtained using MAST, with a fold-change > 2 and adjusted P value < 0.05. (see
680 Methods). The *x*-axis indicates the log2 fold-, and the *y*-axis depicts the significance of the change
681 (-log10 of the adjusted P-value). The color indicates the gene status.
682 **(E)** Gene-term network depicting downregulated genes and corresponding Gene Ontology (GO)
683 molecular functions. GO terms are represented as grey nodes, with size corresponding to the
684 number of associated genes. Genes are colored according to the fold-change values. Enriched GO
685 terms were computed with clusterProfiler.
686 **(F)** Gene-term network illustrating upregulated genes and corresponding molecular functions.
687 Legends are defined in (E).
688 **(G)** Dot plots depicting motif activity of FOX transcription factors in MOGHE-GM and MOGHE-
689 WM neurons. Motif activity was computed using chromVar based on the snATAC-seq datasets.
690 Dot color corresponds to motif activity, and the size corresponds to the proportion of nuclei with
691 open chromatin peaks containing a binding site for the motif. The motif sequence recognized by
692 the regulator is indicated in the left.
693
694
695

SUPPLEMENTARY FIGURE

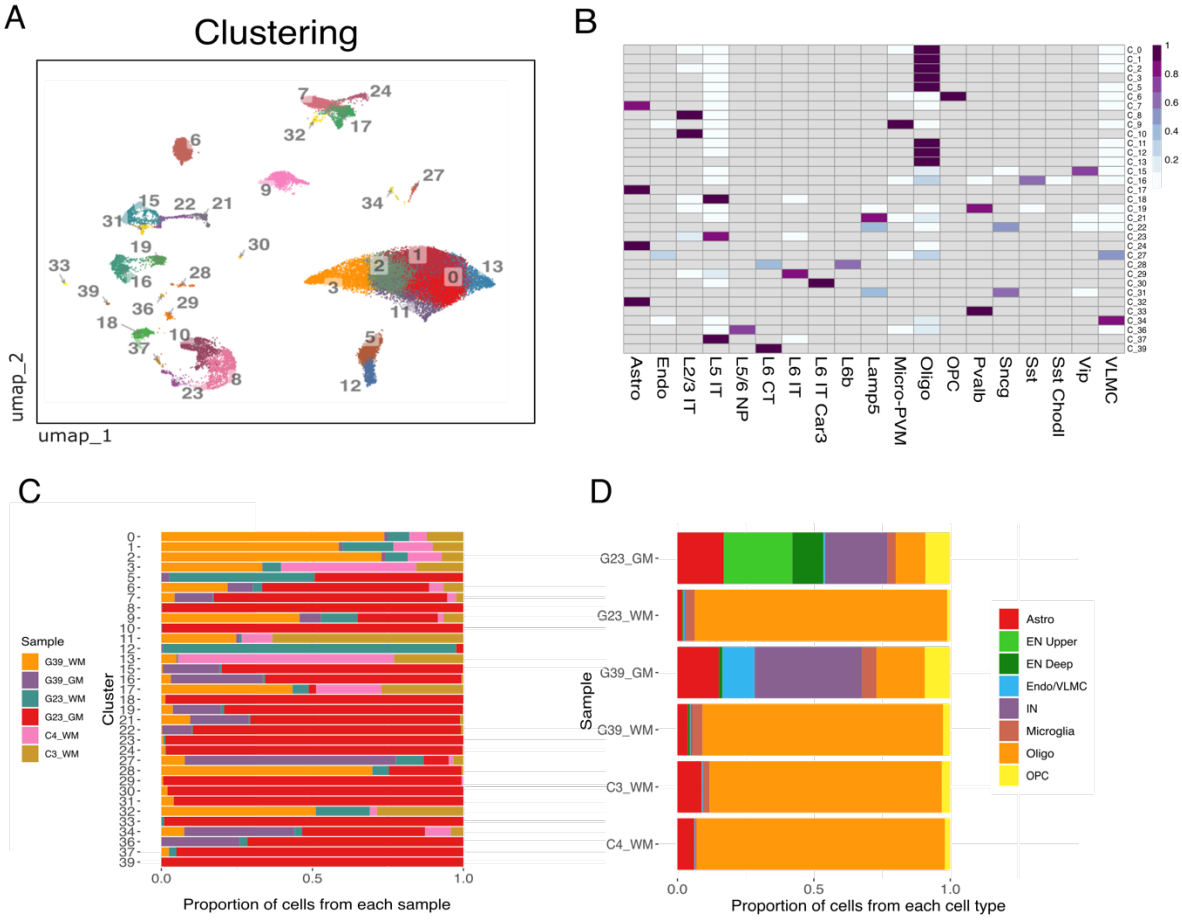


Figure S1, related to Figure 1. Clustering and annotation.

(A) UMAP visualization of nuclei clusters. Clustering was performed in the RNA graph obtained after Harmony integration of MOGHE-WM, MOGHE-GM and Control-WM snRNA-seq samples.

(B) Heatmap depicting the annotation of 40 nuclei clusters performed by Azimuth based on the Allen atlas of the human cortex as a reference. The color indicates the proportion of Azimuth subtypes identified in each cluster.

(C) Bar graphs denoting the representation of samples across nuclei clusters.

(D) Bar graphs showing the proportion of cell types across samples.