Sex differences in progressive multiple sclerosis brain gene expression in oligodendrocytes and OPCs

Brenna A. LaBarre[1,2], Devin King[1,2], Athanasios Ploumakis[2,3], Alfredo Morales Pinzon[4], Charles R.G. Guttmann[4], Nikolaos Patsopoulos[2], *Tanuja Chitnis[1]

Author affiliations:

1. Translational Neuroimmunology Research Center, Ann Romney Center for Neurologic Diseases, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA 02115

2. Systems Biology and Computer Science Program, Ann Romney Center for Neurologic Diseases, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA 02115

3. Spatial Technologies Unit, Beth Israel Deaconess Hospital, Boston, MA, USA 02115

4. Center for Neurological Imaging, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA 02115

Correspondence to: Tanuja Chitnis MD

60 Fenwood Rd, suite 9002K

Hale Building for Transformative Medicine

Boston, MA 02115

tchitnis@bwh.harvard.edu

Running title: MS sex differences in OLs and OPCs

Keywords: Multiple sclerosis; sex-specific; snRNA-seq; oligodendrocytes; OPCs; remyelination

Abstract

Multiple sclerosis is a neurological autoimmune disease with sex-imbalanced incidence; in the USA, the disease is more likely to effect females at a ratio of 3:1. In addition, males are more likely to have a more severe disease course at time of diagnosis. Questions about both causes and downstream effects of this disparity remain. We aim to investigate gene expression differences at a cellular level while considering sex to discover fine-scale sex disparities. These investigations could provide new avenues for treatment targeting, or treatment planning based on sex.

Public single-nuclei RNA-sequencing data from three publications of progressive MS including control brains were analysed using the Seurat R package. Differential gene and pathway expression was looked at both within a specific data set which has sub-lesion level sample dissection and across all studies to provide a broader lens. This allowed for the consideration of cell types and spatial positioning in relation to the interrogated lesion in some of the calculations.

Our analysis showed expression changes in the female MS oligodendrocytes and oligodendrocyte progenitor cells compared to healthy controls, which were not observed in the corresponding male affected cells. Differentially up-regulated genes in females include increased HLA-A in the oligodendrocytes, and increased clusterin in the oligodendrocyte

progenitor cells. There are also several mitochondrial genes in both the oligodendrocytes and oligodendrocyte progenitors which are up-regulated in females, including several directly involved in electron transport and which have previously been associated with neurodegenerative diseases.

These results point to altered states in oligodendrocyte progenitors and oligodendrocytes that in combination with known physiological dissimilarities between sexes may denote different programming in males and females in response to the onset of demyelinating lesions. The potential for increased debris clearance mediated by clusterin and availability of oligodendrocyte progenitors in females may indicate an environment more primed for repair, potentially including remyelination. This could contribute to the disparity in etiology in females versus males.

Introduction

Multiple sclerosis is a common neuroinflammatory autoimmune disease with a sex-biased etiology(1-5). The disease is characterized by the presence of lesions in the brain, and one of the main McDonald diagnostic criteria for the disease is dissemination of these brain lesions in space and time, as found on MRI(6). These lesions are caused by inflammatory damage to myelin along neuronal axons. By utilizing newly developed single nuclei sequencing technologies, researchers are now able to sample cells from plaque-containing post-mortem CNS tissues(7-9) to investigate this cellular environment.

Recently, Jäkel *et al.*(8) describe various sub-groupings of oligodendroglia with altered proportions in MS vs control brains. The authors also observed changes in the normal cell

68 balance amongst samples from normal-appearing tissue adjacent to MS lesions, including a
69 reduction in numbers of oligodendrocyte progenitor cells (OPCs). Work by Schirmer *et al*.(9)
70 highlighted the activated states of many cell types in MS lesions and mapped activated
71 signatures to the rims of chronic active lesions using spatial transcriptomics. In the study by
72 Absinta *et al*.(7), signatures of activated microglia and astrocytes were also observed with
73 unique enrichment in the rims of chronic lesions. One innovation of the Absinta *et al*. study
74 was also that the sampling for the single cell sequencing for these lesions occurred in several
75 places per tissue block, so that there are separate sequencing runs for lesion core, lesion rim,
76 and non-lesion tissue, which allows for some localization of results.

77

78 Disease severity and progression courses in MS vary by sex, with more severe cases being
79 enriched in male subjects. The clinical outcomes of these differences have been investigated
80 and attributed to many causes(2-5, 10), including differences in demyelination and
81 remyelination patterns between sexes. Here, we have combined and compared single-nuclei
82 RNA-seq (snRNA-seq) data from previously published studies to explore sex differences in
83 gene expression across broad cell types. This approach may serve to elucidate molecular
84 mechanisms of these known sex-related differences.

85

86 Three publicly available data sets were downloaded and processed (see Methods)(7-9). As
87 human CNS tissue is not a readily available biospecimen, combining data sets provides us
88 with a larger sample set to interrogate. Our integrated dataset consists of a total of 21 MS
89 subjects (47.6% Female) and 17 Control subjects (35.3% Female), with mean ages between
90 mid-40s and late 50s (Table 1). MS subjects were diagnosed with some form of progressive
91 disease and had mean disease durations of approximately 20 years. Throughout this study,

data from the more recent publication by Absinta *et al.*, were compared to the combined data from Jäkel *et al.* and Schirmer *et al.* to evaluate reproducibility of results.

Table 1. Demographics of included samples

		Publication					
		Absinta *et al.*[7]		**Jäkel *et al.*[8]**		**Schirmer *et al.*[9]**	
		MS	**Control**	**MS**	**Control**	**MS**	**Control**
Sex							
	Female	1	1	1	1	8	4
	Male	4	2	3	4	4	5
Age		50.2 (6.4)	54.7 (5.5)	46.8 (8.4)	58 (17.5)	45.2 (7.0)	53.4 (16.6)
Multiple Sclerosis Type							
	Primary Progressive		-	1	-	1	-
	Secondary Progressive		-	3	-	11	-
Disease duration		19.2 (8.6)		20 (6.6)		19.2 (10.9)	

Ages and disease durations have standard deviations in parentheses. For information on sample selection, causes of death, and IRBs, please see relevant publication. All samples for the Absinta *et al.* MS cohort were indicated to be "progressive", but not whether they were primary or secondary progressive.

Materials and methods

Publicly available data

We utilized data from publicly available snRNA-seq experiments of MS samples for which fastq files were retrieved from SRA. Cell Ranger v6(11) was run to obtain count data, and

then all data sets were loaded into R(12) using Seurat v3(13). Data were collected from Jäkel *et al.*(GSE118257)(8), Schirmer *et al*. (PRJNA544731)(9), and Absinta *et al*. (GSE180759)(7).

Data collection, merging and processing led to a final analyzed data sets consisting of 52,323 cells from the Schirmer *et al*. dataset consisting of 21 individuals, 12 affected and nine controls; 24,890 cells from the Jäkel *et al*. dataset consisting of nine individuals, four affected and five controls; and 58,220 cells from the Absinta *et al*. data set consisting of eight individuals, five affected and three controls.

Data processing

Using Seurat (13), the nuclei were filtered for those that contained information from at least 500 reads, and a minimum of 200 genes but no more than 2500 genes, and those that contained < 20% reads corresponding the mitochondrial genes. Genes which appeared in less than three nuclei were also excluded.

Once cells were filtered, SCTransform(14) was used to normalize the data throughout a given sequencing run. These runs were then merged into a single Seurat object using 2000 variable features for the RunHarmony function from the Harmony(15) R package, which was applied to “harmonize” the data and account for batch effects between sample runs, different sequencing chemistries, and different data sets. To remove “cells” which may be doublets, the R package scDblFinder(16) was used.

Cell type identification

After combining the data, clustering of the nuclei was performed. FindNeighbors was run with 13 dimensions from the harmony dimension reduction, followed by FindClusters run with a resolution of one. A UMAP was then calculated and plotted. Using known marker genes for brain cell types, each numbered cluster was than assigned one of six major cell type labels: Oligodendrocytes (OLs), OPCs, Neurons, Astrocytes, Immune, or Endothelial/Vascular.

Once these initial labels were applied, the cells underwent multiple rounds of relabeling to build confidence in the cell labels. All cells with a given label were subset from the full data, then re-harmonized and re-clustered using the same parameters as were used on the full data. Using the SingleR(17) function in R and the cell type labels provided by the authors of Jäkel *et al.*, a type was assigned to every cell. The more detailed names provided by Jäkel *et al.* were collapsed down to the relevant label for the list of six given above. These labels were then used to filter the clusters. In this first pass filtering, any cluster where 95% or more of the cells were given the same label were sent on for a second round. For those clusters were there were not 95% one label, any label for which there were five or more cells were reserved to be merged later with the matching label, and any label with less than five cells had those cells removed.

The cells for a given label were now subjected to a second round of filtering. Those cells from clusters with 95% the given label, and cells from the other six cell type first round of filtering with the given label that were reserved, were brought back together and again re-merged and re-clustered, using the same parameters as the full data set. For the second filter,

a cluster now had to pass a threshold of 99% the same label to avoid breakup, and any cluster that had less than 99% only had the majority cell type retained, and the rest of the cells were removed. In total, 886 cells are removed with this filtering.

Final cell numbers of twice filtered assigned labels are as follows: OLs – 72,041 cells; OPCs – 6,794; neurons –27,674 cells; astrocytes – 15,862; endothelial/vascular – 4,338; immune – 8,724 cells. For a breakdown of cells by data set, see Supplementary Table 1.

After cell type determination, each of the six major cell types underwent subtype finding. To accomplish this, the clustering step of the Seurat processing per major cell type was re-run; this does not alter the UMAP projection locations, just the cluster assignment of individual cells. A clustering resolution of 0.5 was used which resulted in several subtypes per cell type that were generally supported by the literature. Names were assigned to subtypes based on marker gene lists curated from the literature and/or pathway analyses.

Analyses

Differential expression

Differential expression comparisons were made using both the MS vs Control cell axis as well as cell type and sex specific axes. Using the nebula R package(18) for each comparison (the combination of compartment and sex), only the relevant cells were considered, and model formulas consisted of an intercept and the factor of interest.

Comparison across publications

Analyses were run in three modes: only data from the Absinta *et al*. paper, data from the combined Jäkel *et al.* and Schirmer *et al.* papers, and all three data sets combined. This was done in part because of low sample numbers and to assess reproducibility across data sets. For the sex specific analyses, differential expression was evaluated for each sex in each of these modes. The results from the Absinta *et al.* only analysis were then intersected with the results from the Jäkel *et al.* and Schirmer *et al*. combined analysis, preserving those genes which were statistically significant in both (FDR < 0.05) and shared directionality in the their log fold-change (logFC). This list was then also cross-referenced with the list of differential genes generated by combining all three data sets. With the sex-agnostic analyses a similar protocol was used, with the data being run in three modes, compared between Absinta *et al.* vs Jäkel *et al.* and Schirmer *et al.*, and then checked against all three together.

Pathway analysis

Results from the nebula analyses were then used for pathway enrichment using the fgsea R package(19). FDR values were calculated from the output results of the nebula models and used to determine the input gene lists to the fgsea function. The Canonical Pathways (CP) database from MSigDB from the Broad Institute of MIT and Harvard(20, 21) was used as the query pathways database.

Comparison of expression levels

For comparing expression of specific genes by region, and numbers of cells by region, RNA transcript counts and cell counts were used. In the analysis comparing HLA-A expression and

CD8+ T cells, the number of HLA-A transcripts were counted per compartment, and that number was divided by the number of OLs per compartment, considering only the Absinta *et al* data. The number of cells identified as CD8+ T cells as described in the methods above were also counted, and these values were then plotted in heatmaps. To compare OL and OPC marker gene expression and CLU expression, for each of the genes the transcripts were counted per region and per sex, as were the number of OL and OPC cells; the transcript count was then divided by the corresponding cell count for MBP, PLP1, OLIG1 and NG2 (labeled in the snRNA-seq data as CSPG4). For CLU, transcripts were considered from all cells types, not a specific cell subset, so transcript counts per region were divided by total cells per region, again considering only the Absinta *et al*. data for these analyses.

Visualizations

Results from differential analyses with nebula were visualized with the R package EnhancedVolcano(22). Heatmaps were generated using the ggplot2 R package(23).

Results

Sex differences in gene expression by cell type

Across cell types, we observe differences in gene expression between sexes, when examining sexes individually. In females, there are significantly more differentially expressed genes in the astrocytes and OLs, compared to males, who demonstrate more expression variance in the neurons (Table 2). In the females, many of the genes with differential expression across cell

types are over-expressed in the MS sample compared to control sample (Supplemental Figure 1), whereas in the males the neuronal differential expression is dominated by under-expression of genes compared to the healthy controls (Supplemental Figure 1, Supplemental Figure 2, Figure 1). Those genes which were identified by comparing the data sets were largely recapitulated in the analysis of all three datasets together, and only results which were consistent across data sets were further investigated (Table 2).

Table 2. Significantly differentially expressed genes with coordinated expression between data sets

	Compared between datasets			All data sets together, matching comparison		
Cell	**Female**	**Male**	**Combined**	**Female**	**Male**	**Combined**
Astrocytes	231	1	15	226	1	15
Endothelial	-	-	-	-	-	-
Immune	-	1	6	-	1	6
Neurons	6	40	77	6	40	77
Oligodendrocytes	468	8	27	460	6	27
OPCs	45	-	1	45	-	1

First two columns are the number of genes which were significant (FDR < 0.05) in a differential expression analysis of MS vs Ctrl cells for a given cell type in the Absinta *et al.* data set and were also significant in the same analysis using the combined Jäkel *et al.* and Schirmer *et al.* data set, all having log fold-changes which were the same direction. The last fourth and fifth columns are comparing the results from the first two columns to a sex-specific analysis of all three data sets, to confirm that the results from the "comparison" list weren't artifacts. Columns three and six are the non-sex-specific

Multiple mitochondrial genes, including MT-ATP6, MT-CO2, MT-CYB, MT-ND2, MT-ND4, and MT-ND4L were found to be upregulated in the female OPCs and OLs but not in the corresponding male cells (Figure 1). These genes are important components of the oxidative phosphorylation pathway, pointing to metabolic- or mitochondrial function-related

changes in the female MS cells. While detection of mitochondrial genes can be a marker of lower-quality snRNA-seq data, or an indicator of a high proportion of apoptotic cells (24), our quality control process included steps to mitigate this (see Methods), and these changes are seen across cell types and clusters.

Another notable difference between sexes is the level of change of HLA-A in male versus female OLs. While there is significant increased expression in MS compared to control cells in both males and females, the logFC in females is 3.97 versus 1.05 in males (Figure 1). Looking at the localization of cells and expression in the Absinta *et al*. data set, we find that both the presence of CD8+ T cells and the expression of HLA-A by OLs is highest in the rim portions of the studied lesions in females, with males also having highest HLA-A expression in the rim and similar numbers of CD8+ T cells in the rim and normal-appearing brain tissue (NBT) (Figure 2). There is the least HLA-A expression and fewest CD8+ T cells present in the lesion, and an intermediate amount in the normal-appearing adjacent tissue (there were no NBT samples of female MS subjects in this data set). However, it is important to note that the numbers of cells identified as CD8+ T cells are low across samples, with only 158 cells being considered here.

Sex-agnostic gene expression by cell type

As a second layer of evidence, and to mitigate results which may be due to the small sample sizes inherent when using only the Absinta data set, the full analysis was also done on the data without separating for sex. When performing the analyses combining all the cells of both sexes together, there is a much lower number of differentially expressed genes than in the

sex-specific analysis. The exception here is the neurons, which have a greater number of significant results than in the separated analysis, perhaps due to larger sample size.

Pathway analysis of sex-specific results

Using those genes which were coordinated across studies, a GSEA(19, 21) analysis was run using the canonical pathways (CP) database. In females, we see enrichment for different ribosomal pathways and translational pathways across results for astrocytes, OLs, and OPCs. In the neurons, we see depletion of pathways important for proper neuronal function (Supplemental Figure 3).

In OPCs, a gene contributing to many pathway results is clusterin (CLU). Clusterin was one of the genes that we find to be highly overexpressed in the MS OPCs in females, but unchanged in the male MS OPCs. This gene has previously been implicated in the lack of remyelination in Alzheimer's disease (AD) (25), as well as apoptotic pathways and mitochondrial stabilization(26-29).

To investigate potential function of CLU in these data, the transcription of CLU across lesion compartments was compared to the expression of myelin basic protein (MBP) across compartments in both OLs and OPCs (Figure 3A, 3B, Supplemental Figure 4). MBP is a major component of the intact myelin sheaths of OL cells. Across all MS cells in Absinta *et al*. data, the CLU expression is highest in the rim, while relative MBP expression is highest in the normal-appearing brain tissue for both OLs and OPCs (Supplemental Figure 4). Looking sex-wise, MBP expression does follow a gradient for males from the NBT through the lesion

core (highest -> lowest) in both OLs and OPCs, but in females the strongest MBP expression is in the rim in OLs (8.23 reads per cell in female, vs 6.91 reads per cell in male; results come from 1 female sample and 4 male). This may indicate that MBP in MS lesions is better preserved in females compared to males, either through increased resistance to demyelination or increased propensity for remyelination. Despite clear differences in the MBP expression in OLs, the OPCs show the same pattern in females as in the males. With respect to CLU, both males and females have the highest expression in the rim. Considering the differentiation of OPCs to OLs that would occur during remyelination, we also looked at three more markers, NG2, OLIG1, and PLP1(30, 31)(Figure 3C-E). OLIG1 is expressed at some level across both cell types and throughout differentiation, NG2 is expressed in OPCs and diminishes over differentiation, and PLP1, like MBP is found in mature myelinating OLs. For PLP1, we see a pattern as we would expect, more expression in the more “intact” compartments of OLs, and very low expression across OPCs. NG2 is specific to the OPCs, with the highest levels of expression in the rim cells, which is similar to OLIG1 in the OPCs; in the OLs OLIG1 expression is not as strong.

Among the male DEG results, we see pathways related to ribosomes across cell types, but also pathways related to different types of cellular stress (Supplemental Figure 5) including heat stress and starvation. Especially in the OPCs, there are several pathways related to heat shock factor 1 (HSF1) activation. In the OLs and astrocytes, there is increased expression of ribosome pathways, as well as translation initiation and elongation pathways, suggesting an active cellular response.

For the pathway results which are consistent between the Absinta *et al.* data set and the combined Jäkel *et al.* and Schirmer *et al.* data, there are only recapitulated results in female astrocytes, OLs and OPCs, and in male neurons. Among these results are many pathways related to other neurologic diseases (Supplemental Figure 6).

Discussion

Our results suggest there are sex-specific differences in gene expression and pathway regulation across different cell types in MS brain lesions.

HLA-A implicated in CD8+ T reactivity

HLA-A differential expression in OLs is an interesting result because of the genetic risk factors associated with other HLA antigens, most notably the HLA-DRB1*1501 allele, and the fact that MS is an autoimmune disease (32, 33). Though previously multiple sclerosis has been more strongly associated with MHC-class II antigens (34), HLA-A is a MHC-Class I antigen. A study in HLA-A*0201 transgenic mice which were primed with MOG peptides found a particular peptide (MOG_{181}) that was able to stimulate a strong CD8+ T cytotoxic response; this was also shown to exacerbate MOG_{35-55} induced EAE (33). Similarly, using CD8+ T cells derived from multiple sclerosis affected (and control) subjects, CD8+ T cells were shown to be cytotoxic to HLA-A2 expressing OL cells even without the addition of exogenous MBP peptides (35). Thus, the increased expression of HLA-A that we observed could lead to an increase in CD8+ T cell-mediated OL loss. The higher levels of HLA-A seen in the female OLs suggest that a stronger immune response may be elicited in women.

Clusterin in females could alter OPC differentiation, or aid in debris clearance

Clusterin has previously been investigated in the OPCs of AD patients (25). While this gene is a known risk factor for late-onset AD, it was also found to be upregulated in a subset of non-diseased mouse OPCs; upregulation in OPCs was also seen in our study. The authors found that increased phagocytosis of myelin debris and oligomeric Aβ resulted in increased CLU expression, which in turn inhibited OPC differentiation and new myelin production. Given the previously observed superior OPC activity in female versus male rodent models (4, 10), if CLU is preventing this pool of OPCs from differentiating into remyelinating OLs then this may be a driver of accumulating demyelination and downstream neuronal damage in female multiple sclerosis subjects.

However, the functional potential of CLU is complex, with many characterized and competing roles attributed to it. There exists a nuclear-located isoform promoting apoptosis(27), a mitochondrial form that averts apoptosis by preventing mitochondrial membrane permeabilization(29), and an excreted form that mainly serves as an extracellular chaperone for misfolded proteins (28). Data used here are only from mRNA, and while there are transcriptional differences between some of these isoforms which may be present in the sequencing data, those analyses were outside the scope of this study.

Considering the comparison of MBP expression and CLU expression in OLs and OPCs, the heightened expression of MBP in the female OL rim may indicate better preservation of the

myelin in this region, either linked to increased resistance to demyelination or increased propensity to remyelinate. If the CLU expression in these samples indicates an apoptotic function, the rim would be where one would expect to see increased cell loss. However, studies in mice after brain ischemia showed mice which overexpressed CLU had better recovery than wild type and CLU -/- mice(36), which may indicate a high expression of CLU is beneficial to debris clearage in its role as an extracellular chaperone and may be anti-inflammatory and conducive to remyelination.

These results, in combination with the high expression of NG2 in the rim OPCs, might also indicate a stalling of remyelination at this site. As NG2 expression in OPCs goes down in the lesion compartment, this may indicate differentiation of these cells. The cells in the rim, however, maybe be pausing or accumulating before repairing the lesion damage. This could again be attributed to the CLU expression; potentially in two functions. The prevention of differentiation as seen in Beiter *et al.(25)* and aiding debris clearance, which may work in tandem to produce better lesion recovery over all.

These differences in OPC and OL capacity and functionality may contribute to the differences in disease course seen across sexes; females are able to recover from attacks with their greater pool of active OPCs for some time until they are disabled, whereas males have a higher initial proportion of OLs but less regenerative capacity and once they have reached critical OL loss enter a clinically recognizable and primarily progressive disease phase (3).

Mitochondrial components are upregulated in female OPCs and OLs

Several of the pathways which were found to be recurrently dysregulated across cell types, sexes, and studies implicate oxidative phosphorylation and other mitochondrial functions. There is growing literature about the implications of mitochondrial dysfunction in multiple sclerosis. In progressive EAE models it was found that axonal mitochondrial homeostasis was disrupted prior to the symptom onset stage(37). It has also been shown that changes in mitochondrial function correlated with lesion disease course and neurological functions (38). Also, mitochondrial collapse due to the loss of the mitochondrial transmembrane potential can trigger cell death(39).

However, while altered mitochondrial function could lead to cell loss, there is also evidence that increased expression of mitochondrial genes can be triggered by the differentiation of OPCs(40), which in our data would indicate that the female OPCs are better able to differentiate than the male OPCs, and therefore lead to better recovery.

Limitations

For this study, we used data sets produced by other laboratories that were made publicly available. While this data is a useful resource, it is a very small sample set, and in this case most of the major findings were initially identified in only one female multiple sclerosis sample (though they were corroborated with a larger data set with more female multiple sclerosis samples). Also, the snRNA-seq quality for these experiments, while sufficient for the method at the time of experiments, could now be improved upon. There is some

394 possibility for cross-cell contamination which may affect the ability to assign a cell type to
395 each individual sequenced "cell". While this should be accounted for in many of the data
396 processing methods that we employed in quality control steps, it is unlikely for all
397 confounding noise to be eliminated.

398

399 Also, given that the nature of this study is a reanalysis of multiple data sets in a difficult to
400 obtain, post-mortem human tissue, in vivo and/or in vitro validation of the results presented
401 here are outside of the scope of this project; these results do however offer strong candidates
402 for further follow-up experiments.

403

404 # Conclusions

405 Though this was an exploratory study based on public data, we were able to investigate sex
406 differences in cell-specific subsets of MS brain tissue. Our results indicate that there are
407 alterations in gene expression patterns in MS, which vary by sex and cell type. Especially in
408 the OPCs and OLs, we see a much greater change in expression patterns in the female.

409 These results point to a potentially more regenerative pattern in the female compared to the
410 male brain, leading to more aggressive clinical progression phenotypes in males and a
411 delayed decrease in regenerative ability in females over the course of disease.

Declarations

Ethics approval and consent to participate

Samples and data used in this study were collected as approved in the referenced studies. IRB approval was not required for the analysis presented here.

Consent for Publication

Not applicable

Data availability

Data were collected from Jäkel *et al.*(GSE118257)(8), Schirmer *et al.* (PRJNA544731)(9), and Absinta *et al.* (GSE180759)(7) through the resources provided by the National Center for Biotechnology Information (NCBI) through the National Library of Medicine at the National Institutes of Health, US.

Funding

This work was funded by the National Institutes of Health grant R21NS123826 (to TC).

Competing interests

Authors declare they have no competing interests. NP is now an employee of Novartis.

Author Contributions

Authors contributed the following: BAL, NP, CRGG, TC designed the study. AP, AMP, DK collected data. BAL, DK, processed and analysed data. BAL and TC produced manuscript and figures. All authors reviewed, edited and approved the final manuscript.

Acknowledgements

Thank you to Prashanth Sama for technical advice.

References

1. Smith-Bouvier DL, Divekar AA, Sasidhar M, Du S, Tiwari-Woodruff SK, King JK, et al. A role for sex chromosome complement in the female bias in autoimmune disease. J Exp Med. 2008;205(5):1099-108.

2. Voskuhl RR, Pitchekian-Halabi H, MacKenzie-Graham A, McFarland HF, Raine CS. Gender differences in autoimmune demyelination in the mouse: implications for multiple sclerosis. Ann Neurol. 1996;39(6):724-33.

3. Cerghet M, Skoff RP, Bessert D, Zhang Z, Mullins C, Ghandour MS. Proliferation and death of oligodendrocytes and myelin proteins are differentially regulated in male and female rodents. J Neurosci. 2006;26(5):1439-47.

4. Li WW, Penderis J, Zhao C, Schumacher M, Franklin RJ. Females remyelinate more efficiently than males following demyelination in the aged but not young adult CNS. Exp Neurol. 2006;202(1):250-4.

5. Alvarez-Sanchez N, Dunn SE. Potential biological contributers to the sex difference in multiple sclerosis progression. Front Immunol. 2023;14:1175874.

6. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018;17(2):162-73.

7. Absinta M, Maric D, Gharagozloo M, Garton T, Smith MD, Jin J, et al. A lymphocyte-microglia-astrocyte axis in chronic active multiple sclerosis. Nature. 2021;597(7878):709-14.

8. Jäkel S, Agirre E, Mendanha Falcao A, van Bruggen D, Lee KW, Knuesel I, et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. Nature. 2019;566(7745):543-7.

9. Schirmer L, Velmeshev D, Holmqvist S, Kaufmann M, Werneburg S, Jung D, et al. Neuronal vulnerability and multilineage diversity in multiple sclerosis. Nature. 2019;573(7772):75-82.

10. Yasuda K, Maki T, Kinoshita H, Kaji S, Toyokawa M, Nishigori R, et al. Sex-specific differences in transcriptomic profiles and cellular characteristics of oligodendrocyte precursor cells. Stem Cell Res. 2020;46:101866.

11. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8:14049.

12. Team RC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2020.

13. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888-902 e21.

14. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20(1):296.

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

16. Germain P-L, Lun A, Macnair W, Robinson MD. Doublet identification in single-cell sequencing data using scDblFinder. F1000Research. 2021;10.

17. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol. 2019;20(2):163-72.

18. He L, Davila-Velderrain J, Sumida TS, Hafler DA, Kellis M, Kulminski AM. NEBULA is a fast negative binomial mixed model for differential or co-expression analysis of large-scale multi-subject single-cell data. Commun Biol. 2021;4(1):629.

19. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. 2021.

20. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27(12):1739-40.

21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

22. Blighe KR, Sharmila; Turkes, Emir; Ostendorf, Benjamin; Grioni, Andrea; Lewis, Myles EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. 1.18.0 ed2023. p. R package.

493 23. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New
494 York; 2016.

495 24. 10X Genomics. Why do I see a high level of mitochondrial gene expression? 2018
496 [Available from: https://kb.10xgenomics.com/hc/en-us/articles/360001086611-Why-do-I-see-
497 a-high-level-of-mitochondrial-gene-expression-.

498 25. Beiter RM, Ennerfelt HE, Rivet-Noor C, Merchak A, Bai R, Johanson DM, et al.
499 Oligomeric amyloid beta prevents myelination in a clusterin dependent manner. biorXiv.
500 2022.

501 26. Herring SK, Moon HJ, Rawal P, Chhibber A, Zhao L. Brain clusterin protein isoforms
502 and mitochondrial localization. Elife. 2019;8.

503 27. Leskov KS, Klokov DY, Li J, Kinsella TJ, Boothman DA. Synthesis and functional
504 analyses of nuclear clusterin, a cell death protein. J Biol Chem. 2003;278(13):11590-600.

505 28. Wyatt AR, Yerbury JJ, Berghofer P, Greguric I, Katsifis A, Dobson CM, Wilson MR.
506 Clusterin facilitates in vivo clearance of extracellular misfolded proteins. Cell Mol Life Sci.
507 2011;68(23):3919-31.

508 29. Zhang H, Kim JK, Edwards CA, Xu Z, Taichman R, Wang CY. Clusterin inhibits
509 apoptosis by interacting with activated Bax. Nat Cell Biol. 2005;7(9):909-15.

510 30. Woodruff RH, Tekki-Kessaris N, Stiles CD, Rowitch DH, Richardson WD.
511 Oligodendrocyte development in the spinal cord and telencephalon: common themes and new
512 perspectives. Int J Dev Neurosci. 2001;19(4):379-85.

513 31. Seeker LA, Williams A. Oligodendroglia heterogeneity in the human central nervous
514 system. Acta Neuropathol. 2022;143(2):143-57.

515 32. Bruijstens AL, Wong YYM, van Pelt DE, van der Linden PJE, Haasnoot GW,
516 Hintzen RQ, et al. HLA association in MOG-IgG- and AQP4-IgG-related disorders of the
517 CNS in the Dutch population. Neurol Neuroimmunol Neuroinflamm. 2020;7(3).

518 33. Mars LT, Bauer J, Gross DA, Bucciarelli F, Firat H, Hudrisier D, et al. CD8 T cell
519 responses to myelin oligodendrocyte glycoprotein-derived peptides in humanized HLA-
520 A*0201-transgenic mice. J Immunol. 2007;179(8):5090-8.

521 34. Ramagopalan SV, Ebers GC. Multiple sclerosis: major histocompatibility complexity
522 and antigen presentation. Genome Med. 2009;1(11):105.

523 35. Jurewicz A, Biddison WE, Antel JP. MHC Class I-Restricted Lysis of Human
524 Oligodendrocytes by Myelin Basic Protein Peptide-Specific CD8 T Lymphocytes. The
525 Journal of Immunology. 1998;160(6):3056-9.

526 36. Wehrli P, Charnay Y, Vallet P, Zhu G, Harmony J, Aronow B, et al. Inhibition of
527 post-ischemic brain injury by clusterin overexpression. Nat Med. 2001;7(9):977-9.

528 37. Buonvicino D, Ranieri G, Guasti D, Pistolesi A, La Rocca AI, Rapizzi E, Chiarugi A.
529 Early derangement of axonal mitochondria occurs in a mouse model of progressive but not
530 relapsing-remitting multiple sclerosis. Neurobiol Dis. 2023;178:106015.

531 38. Sadeghian M, Mastrolia V, Rezaei Haddad A, Mosley A, Mullali G, Schiza D, et al.
532 Mitochondrial dysfunction is an important cause of neurological deficits in an inflammatory
533 model of multiple sclerosis. Sci Rep. 2016;6:33249.

534 39. Bonora M, Wieckowski MR, Chinopoulos C, Kepp O, Kroemer G, Galluzzi L, Pinton
535 P. Molecular mechanisms of cell death: central implication of ATP synthase in mitochondrial
536 permeability transition. Oncogene. 2015;34(12):1475-86.

537 40. Schoenfeld R, Wong A, Silva J, Li M, Itoh A, Horiuchi M, et al. Oligodendroglial
538 differentiation induces mitochondrial genes and inhibition of mitochondrial function
539 represses oligodendroglial differentiation. Mitochondrion. 2010;10(2):143-50.

540

541 Figure legends

542 **Figure 1. Genes which are differentially expressed in each cell subset.** Genes represented
543 are all those that were significant in male Oligos and female OPCs; this data is a subset of the
544 data shown in the heatmap in Supplemental Figure 1; includes all genes which were
545 significant in male OLs and female OPCs

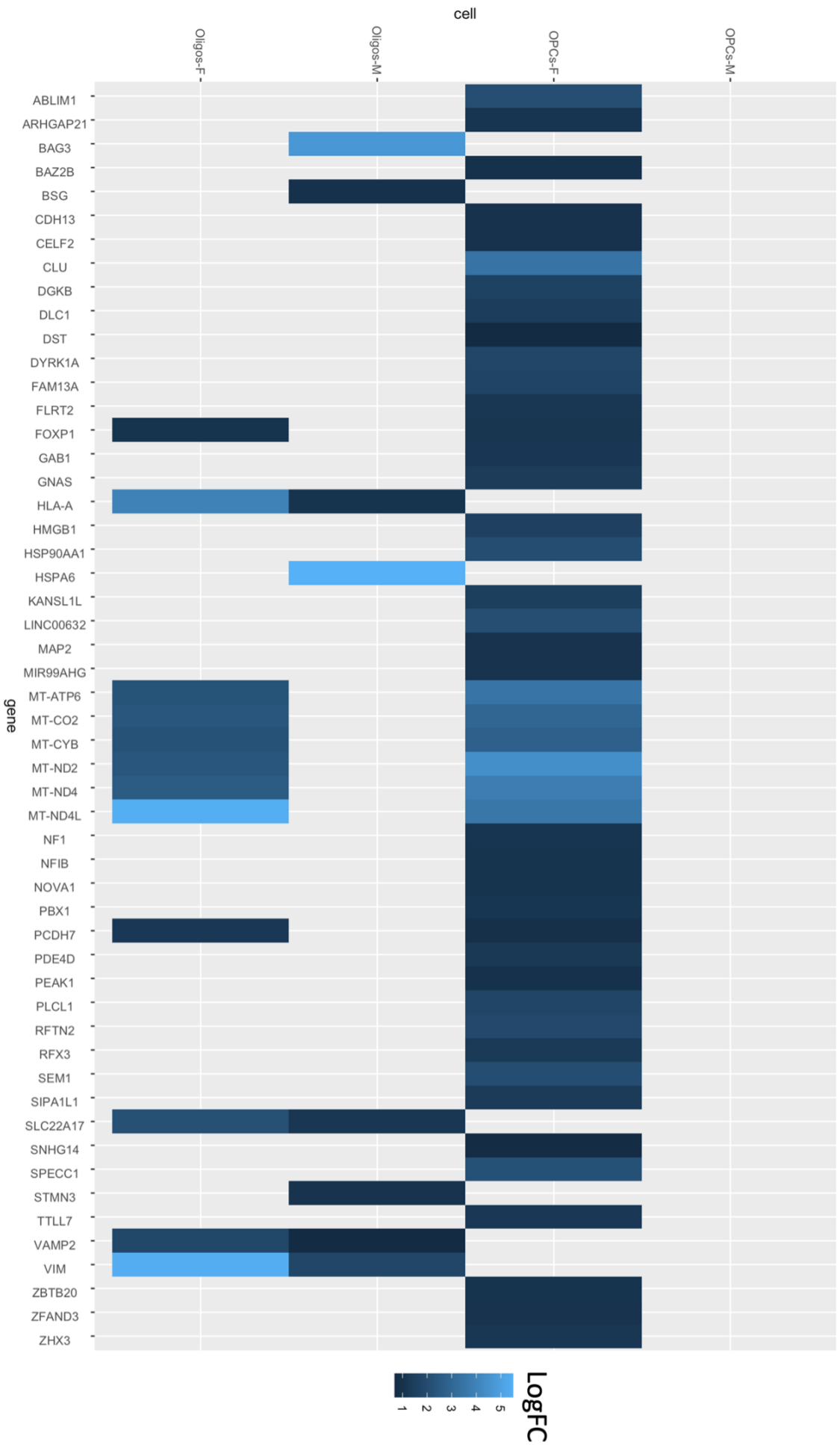
546

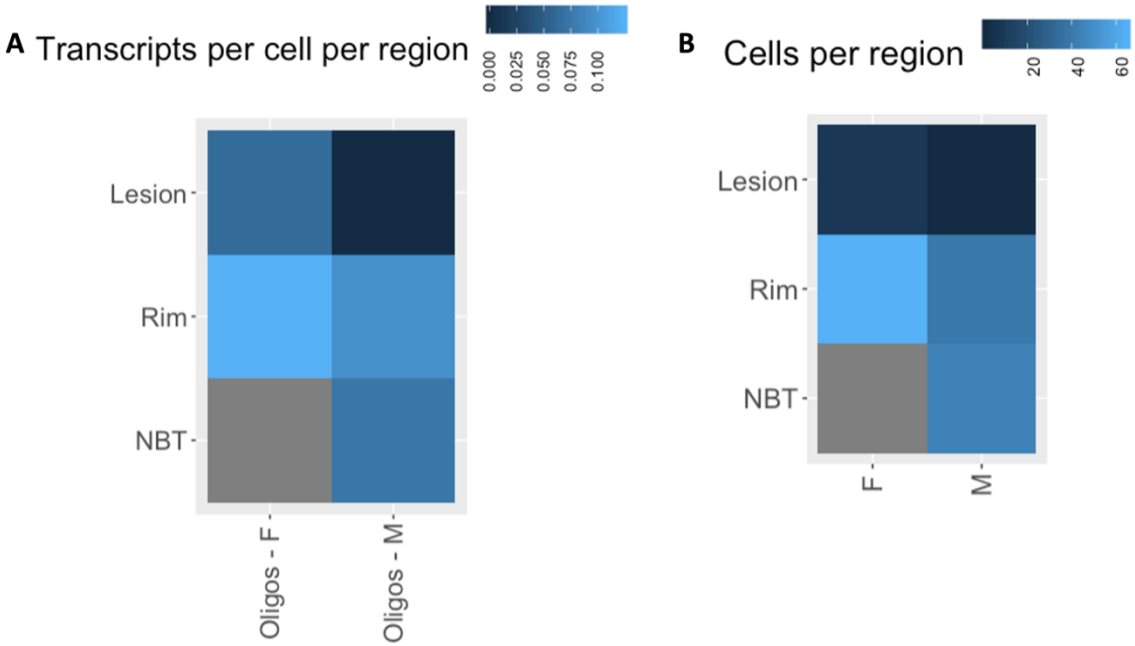
547 **Figure 2. Heatmap of HLA-A transcription in OLs and CD8+ T localization by region**
548 **examined in Absinta *et al*. data.** For the OLs, transcripts and cells for each region were
549 counted, and plotted values are count of transcripts divided by number of cells for each
550 region. For CD8+ T cells, plotted values indicate number of cells identified in each region.
551 Dark blue indicates lower values, light blue indicates higher values. NBT – normal-appearing
552 brain tissue

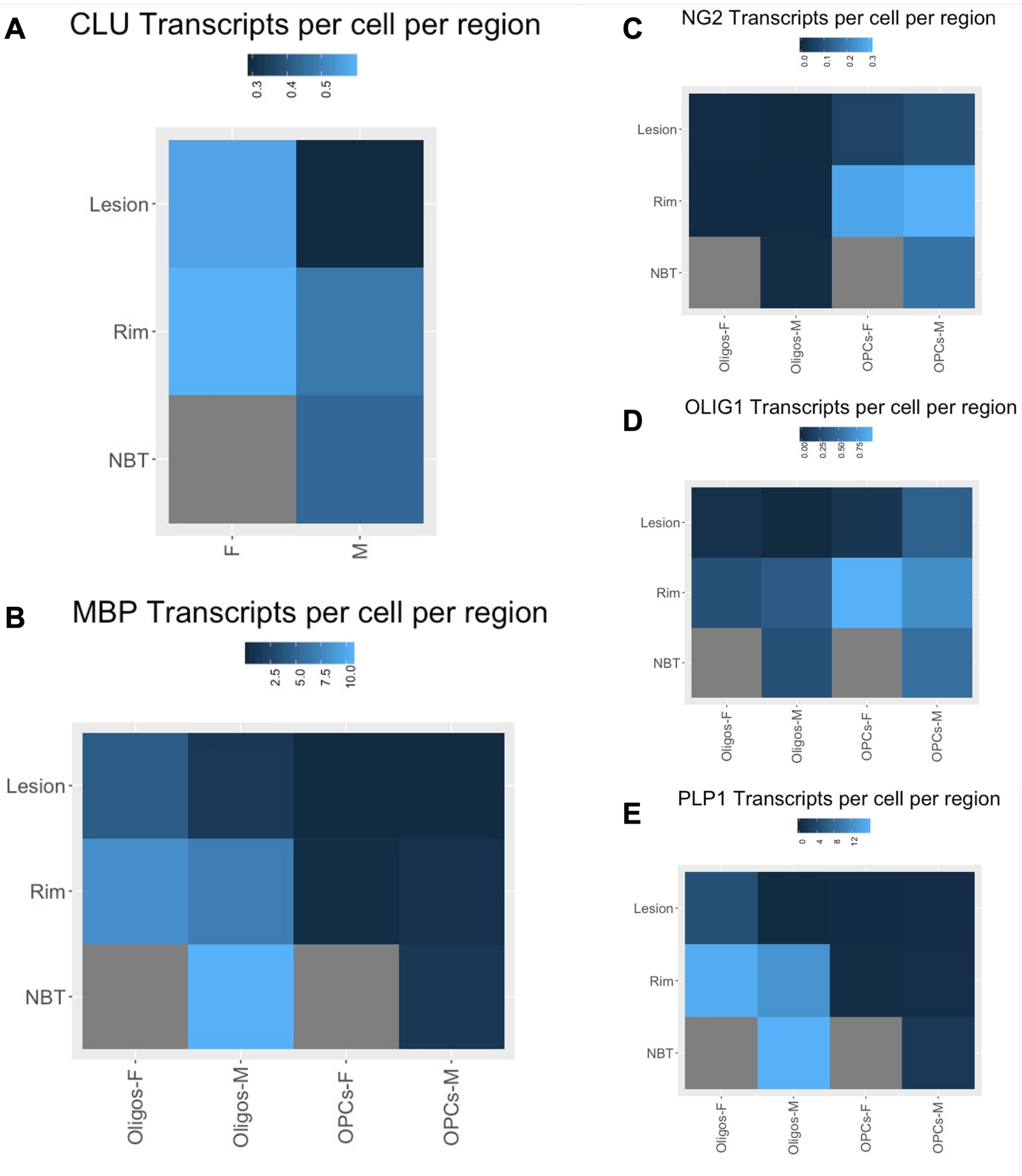
553

554 **Figure 3. Relative expression by sex of MBP, PLP1, OLIG1 and NG2 in OLs and OPCs**
555 **across lesion regions compared to CLU expression from all cells across lesion regions in**
556 **Absinta *et al*. data.** NBT – normal-appearing brain tissue

557


cell
Oligos-F
Oligos-M
OPCs-F
OPCs-M
gene
ABLIM1
ARHGAP21
BAG3
BAZ2B
BSG
CDH13
CELF2
CLU
DGKB
DLC1
DST
DYRK1A
FAM13A
FLRT2
FOXP1
GAB1
GNAS
HLA-A
HMGB1
HSP90AA1
HSPA6
KANSL1L
LINC00632
MAP2
MIR99AHG
MT-ATP6
MT-CO2
MT-CYB
MT-ND2
MT-ND4
MT-ND4L
NF1
NFIB
NOVA1
PBX1
PCDH7
PDE4D
PEAK1
PLCL1
RFTN2
RFX3
SEM1
SIPA1L1
SLC22A17
SNHG14
SPECC1
STMN3
TTLL7
VAMP2
VIM
ZBTB20
ZFAND3
ZHX3
1
2
3
4
5
LogFC


A Transcripts per cell per region
0.000
0.025
0.050
0.075
0.100
Lesion
Rim
NBT
Oligos - F
Oligos - M
B Cells per region
20
40
60
Lesion
Rim
NBT
F
M


A
CLU Transcripts per cell per region
0.3
0.4
0.5
Lesion
Rim
NBT
F
M
B
MBP Transcripts per cell per region
2.5
5.0
7.5
10.0
Lesion
Rim
NBT
Oligos-F
Oligos-M
OPCs-F
OPCs-M
C
NG2 Transcripts per cell per region
0.0
0.1
0.2
0.3
Lesion
Rim
NBT
Oligos-F
Oligos-M
OPCs-F
OPCs-M
D
OLIG1 Transcripts per cell per region
0.00
0.25
0.50
0.75
Lesion
Rim
NBT
Oligos-F
Oligos-M
OPCs-F
OPCs-M
E
PLP1 Transcripts per cell per region
0
4
8
12
Lesion
Rim
NBT
Oligos-F
Oligos-M
OPCs-F
OPCs-M