A comparative transcriptomic analysis of mouse demyelination models and Multiple Sclerosis lesions

Authors:

Erin L Aboelnour[1], Veronica R Vanoverbeke[1], Madelyn M Hatfield[1], and Katrina L Adams[1,2]

Affiliations:

[1]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA.

[2]Center for Stem Cells and Regenerative Medicine, University of Notre Dame, Notre Dame, IN, USA.

Abstract:

Demyelinating diseases, such as Multiple Sclerosis (MS), are debilitating conditions characterized by loss of the myelin sheaths, ultimately leading to neurodegeneration. Toxicity models are among the most commonly used mouse models to induce demyelination; however, it remains unclear whether different demyelination models elicit distinct glial responses, and how comparable these changes are to MS. To address this gap, we integrated new and published single cell transcriptomic data of the subcortical white matter from lysophosphatidylcholine (LPC) and cuprizone toxicity models, and compared them to an existing human MS dataset. We find that LPC and cuprizone treatments induce distinct oligodendrocyte (OL) states, but a highly conserved microglial response upon demyelination. Interestingly, remyelinating OLs converge on an altered maturation state in both LPC and cuprizone models, potentially due to persistent activation of microglia at remyelination stages. Comparison of the mouse models with MS tissue reveals that key OL gene signatures specific to LPC and cuprizone demyelination are observed in MS patients, while microglia appear more heterogeneous across the different types of MS lesions. Finally, cross-species analysis highlights a conserved phenotype shared between cuprizone and actively demyelinating MS lesions, with downregulation of genes required for stable myelin production and increased cellular stress pathways. Overall, this comparative analysis uncovers specific gene expression differences between mouse demyelination models and human MS lesions, providing a foundation for using the animal models effectively to advance remyelination therapies.

Highlights:

- LPC and cuprizone demyelination induce distinct disease-associated oligodendrocyte states with shared pathological changes to MS lesions.
- Remyelinating oligodendrocytes converge on an altered maturation state.
- Microglia response to demyelination is highly conserved in mice but comparatively more heterogeneous in humans.
- Cuprizone treatment induces a disease-associated oligodendrocyte state resembling MS lesion pathology, with downregulation of genes required to produce stable myelin.

Introduction:

Demyelination, the loss of myelin sheaths around neuronal axons, is a major contributor to neurodegenerative disease, such as Multiple Sclerosis (MS). Demyelination is spontaneously repaired in the young, healthy central nervous system (CNS), and is sufficient to provide neuroprotection and slow disease progression [1,2]. However, evidence of remyelination is highly variable across MS patients, and there are no approved treatments that promote remyelination. Therefore, there is an urgent need to understand the biology underlying remyelination. A critical gap in the current use of animal models to study remyelination is the lack of understanding of which MS disease mechanisms are replicated. Remyelination is a complex, multicellular process that is most frequently studied using toxicity-induced demyelination in mice, due to the predictable timeline of damage and repair. The most widely used toxicity models to induce demyelination use LPC or cuprizone. LPC induces a highly tractable lesion at the site of injection, whereas dietary consumption of cuprizone induces OL loss more broadly and in multiple brain structures [3,4]. LPC is a detergent that creates a focal lesion that undergoes repair through reproducible phases of demyelination, increased reactive gliosis, and differentiation of oligodendrocyte precursor cells (OPCs) into OLs within approximately four weeks. Cuprizone is a copper chelator that, when fed to mice for four to six weeks, induces OL cell death throughout the gray and white matter (GM and WM, respectively) of the brain. Then, when mice are returned to their normal food, the brain undergoes spontaneous OPC differentiation and remyelination within two to three weeks [5,6]. These two toxicity models are often used interchangeably by researchers, with a particular method selected because of experimental design constraints rather than scientific rationale.

However, it remains unclear whether LPC and cuprizone induce similar molecular changes and cellular states.

The advancement of single cell and single nuclei RNA sequencing (sc/snRNA-seq) and supporting bioinformatic tools has enabled unbiased characterization of the molecular response across neuronal and glial populations from animal models of demyelination and post-mortem tissue samples from human MS patients. Previous studies have characterized the response to LPC and cuprizone treatment using scRNA-seq [7–9]. In particular, Pandey et al. described three distinct disease-associated oligodendrocyte (DAO) states in the mouse demyelination models that are similar in gene expression to those found in mouse models of Alzheimer's disease; however, they did not directly compare the effects of LPC and cuprizone treatments [8]. Likewise, snRNA-seq analysis of post-mortem human brain samples has provided valuable insight into the cellular and molecular pathology of MS [10–13]. MS lesions, particularly the rim of chronic active lesions, contain stressed OLs, reactive astrocytes, and activated phagocytosing microglia/macrophages [10,14,15]. While mouse models have been used to identify altered signaling pathways that appear conserved in MS lesions, a direct comparison of mouse toxicity-induced demyelination models to human MS lesions using single cell transcriptomics across multiple cell types has not been performed [8,16]. A greater understanding of the similarities and differences between mouse models and MS lesions is critical for effectively using these models to identify and test new therapeutic targets.

In this study, we directly compared LPC and cuprizone models of demyelination, focusing solely on the subcortical WM of the brain. We integrated new and existing sc/snRNA-seq datasets of the demyelinated and remyelinated corpus callosum from LPC- and cuprizone-treated mice [7,8]. In addition, we reanalyzed a publicly available snRNA-seq dataset of the largest cohort of MS lesions to date [13]. We found that LPC and cuprizone induced distinct DAO states, and that both models shared some transcriptional similarities with DAOs in MS lesions. Interestingly, we found the opposite with microglia – LPC and cuprizone-induced demyelination elicited a highly conserved response in microglia, with slightly altered recovery dynamics. However, microglia in human MS lesions displayed very distinct cell states across different lesion types, which were not well recapitulated in the mouse models. Finally, we performed a cross-species comparison of the mouse and human WM, and found that cuprizone, but not LPC demyelination, induced a DAO state that resembled the human DAO state in active MS lesions.

Results

Conserved glial responses in the white matter of mouse demyelination models and MS patient samples.

To compare the effects of LPC- and cuprizone-induced demyelination, we performed snRNA-seq of the adult mouse corpus callosum, the largest WM structure in the mammalian brain. We integrated these data with published scRNA-seq studies of the corpus callosum following LPC- and cuprizone-induced demyelination [7,8]. Together, these datasets cover a time course from baseline (homeostatic conditions) through demyelination and subsequent remyelination (Supplementary Data 1). We obtained 112,017 high-quality filtered cells/nuclei, and the fully integrated UMAP displayed strong overlap between datasets and transcriptome types while maintaining clear cell-type separation (Figure 1A, Supplementary Figure 1A-C). Using expression of known marker genes, we identified clusters corresponding to OL lineage cells, microglia, border-associated macrophages (BAMs), astrocytes, neural progenitor cells (NPCs), ependymal cells, endothelial cells, vascular cells, pericytes, T cells, and three clusters undergoing cell division (OPCs, microglia, and NPCs) (Figure 1B, Supplementary Figure 1D, E).

Changes in cell type abundances were evident in both LPC and cuprizone treatment. As expected, we observed a reduction in mature oligodendrocytes (MOLs), and an increase in microglia populations (Figure 1C, D, Supplementary Figure 1F, G). Although astrocytes are known to respond robustly in both demyelination models, we did not obtain sufficient numbers of astrocytes for robust downstream analysis [17,18]. Interestingly, both MOLs and microglia, which comprised the majority of sequenced cells, underwent transcriptional state shifts in the treated conditions, and the changes in microglia persisted through remyelination. We also observed that BAMs and T cells were specifically enriched in the LPC model, showing distinctions at the cellular level between LPC- and cuprizone-induced response.

To compare the mouse demyelination datasets with human MS lesions, we integrated WM samples from a recently published snRNA-seq atlas of control and MS patient brain tissue [13]. We obtained 321,565 quality-filtered single nuclei, and produced a new embedding and annotation that was similar to the published dataset (Figure 1E, Supplementary Figure 2A-C). As in the mouse dataset, the human dataset was composed primarily of MOLs, microglia, and astrocytes (Figure 1F, Supplementary Figure 2D-E). A major advantage of this dataset is that it was carefully

annotated by lesion type using immunohistochemistry to define active lesions (ALs), chronic active lesions (CALs), chronic inactive lesions (CILs), and remyelinated lesions (RLs) [13]. The WM-only dataset includes 13 control WM samples from non-MS patients, 17 normal-appearing white matter (NAWM) samples, 21 ALs, 17 CALs, 13 CILs, and 11 RLs (Supplementary Data 2, Supplementary Figure 2F).

We then plotted UMAPs by MS lesion type, which revealed shifts in MOL, microglia and astrocyte states across distinct lesion types, mirroring our observations in the mouse demyelination models (Figure 1G). We observed a modest decrease in the proportion of MOLs, with increases in microglia and astrocytes, across the different MS lesions, with ALs potentially exhibiting the largest shifts (Figure 1H). Peripheral immune cells, including T and B cells, were predominantly detected in MS lesion samples, consistent with their known pathological roles in MS. Together, these findings suggest that mouse demyelination models share key similarities at a cellular level with human MS pathology, allowing for comparative analyses across species.

Mouse oligodendrocyte lineage cells exhibit distinct states in LPC and cuprizone treatment.

To understand how specific glial populations respond to the demyelination, we focused on individual cell types. First, we subclustered and reannotated the mouse OL lineage cells using known markers (Figure 2A, Supplementary Figure 3A). We observed a clear developmental progression of progenitors into mature OL (MOL) states: OPCs sequentially become committed oligodendrocyte progenitors (COPs), newly formed oligodendrocytes (NFOLs), and myelin-forming oligodendrocytes (MFOLs). We also defined four MOL states (MOL_A-D). Interestingly, connectivity analysis showed that MOL states were linked by a main trajectory from MOL_A to MOL_B, C, and D, with additional branch points leading to three subclusters we defined as DAO states (MOL_E-G) (Figure 2A) [19]. We used marker genes defined by previous studies to delineate the MOLs, though the functionality of such states remains unknown (Figure 2B, Supplementary Figure 3B, C) [20]. We defined MOLs by the combined expression of mature markers (*Plp1*, *Mog*, *Mag*) and cluster-specific markers: MOL_A: *Ptgds*, *Opalin*, low *Il33*; MOL_B: increased expression of *Il33*; MOL_C: increased expression of *Selenop*, *Spock3*, *Anln*; and MOL_D: expression of *Klk6*, *S100b*, *Anxa5* . MOL_B and MOL_C constituted the majority of corpus callosum OLs, accounting for 75-80% of MOLs at baseline, consistent with a previous report of low *Klk6* expression in the adult mouse corpus callosum [21]. Additionally, we identified the

expression of genes associated with previously described pathological states in the DAO clusters, including *Serpina3n*, *C4b* and *H2-D1* (MOL_E and F) and *Cdkn1a*, *Ddit3* and *Nupr1* (MOL_G) [8,9,16,22].

We hypothesized that the DAO states would be enriched in the demyelination treatment samples. To test this, we generated density plots to visualize the cell distribution across conditions. We saw accumulations of distinct MOL states in LPC- and cuprizone-treated models (Figure 2C). Baseline samples contained predominantly MOL_A–D, with minimal progenitors (OPCs, COPs, NFOLs and MFOLs) or DAOs. LPC-induced demyelination increased progenitor populations, reduced MOL_A-D, and increased DAO clusters MOL_E and MOL_F. In contrast, cuprizone-induced demyelination led to a loss of MOL_A-D and a dramatic shift to the DAO state MOL_G. At remyelination time points, both LPC- and cuprizone-treated samples were enriched in the DAO state MOL_E, suggesting both models converged on a similar remyelination state (Supplementary Figure 3D, E). To quantify these changes, we calculated the differential abundance (DA) across overlapping cell neighborhoods (nhoods) using Milo (Dann et al. 2022). As expected, both treatments resulted in widespread MOL loss and moderate gains in progenitor cells at demyelination (Figure 2D, Supplementary Figure 3F). However, LPC-induced demyelination preferentially enriched DAO states MOL_E and F, whereas cuprizone-induced demyelination resulted in an accumulation of the DAO MOL_G state, confirming the density plot findings. Interestingly, LPC-treated samples at remyelination showed near-complete recovery of homeostatic MOL populations, whereas cuprizone-treated samples exhibited persistent MOL depletion, with a notable increase in MOL_B. This discrepancy may reflect slight differences in remyelination time points between the two models. However, both models enriched the DAO cluster MOL_E at remyelination, suggesting a shared remyelination state. Notably, MFOLs showed a direct connection to MOL_E, suggesting that OLs may undergo an altered maturation state following demyelination (Figure 2A). Together, these results reveal that LPC and cuprizone induce distinct demyelination-associated MOL states that converge on a similar remyelination pathway.

LPC and cuprizone treatment induce transcriptionally distinct disease-associated OLs at demyelination.

We next asked how the different DAO states compare to each other transcriptionally. We tested for differentially expressed genes (DEGs) using a pseudobulk approach (see Methods,

Supplementary Data 3, Supplementary Figure 4A). We found that LPC-induced MOL_E and MOL_F share a common gene signature, whereas cuprizone MOL_G has a largely distinct set of DEGs (Figure 2E, Supplementary Figure 4B). To characterize these differences, we performed gene set enrichment analysis (GSEA) to identify pathways specific to each state (Figure 2F) [23]. TNFa signaling via NFκB was the only pathway enriched across all DAO states, supported by shared upregulation of the genes *Csf1, Fos, Phlda1,* and *Plpp3*, which may contribute to apoptosis activation and p53-mediated inflammation. LPC-associated DAO clusters (MOL_E and MOL_F) are also enriched for interferon gamma (IFNγ) and interferon alpha (IFNα) signaling, marked by upregulation of *B2m*, *Irf9*, *Ifi27*, *H2-D1, Stat1/3*, and *Socs3.* IFN signaling has been linked to disease-associated glia in the experimental autoimmune encephalomyelitis (EAE) model, though it is notably not upregulated at the demyelination timepoint in the cuprizone model [16]. Additionally, coagulation-related genes were enriched in LPC-treated DAOs, likely due to blood-brain barrier disruption and BAM infiltration following stereotaxic injection (Figure 1C, D). In contrast, DEGs in the cuprizone-specific DAO state (MOL_G) were enriched for Myc targets, p53 response, oxidative phosphorylation, and unfolded protein response (UPR), indicating increased protein synthesis and dysregulated stress responses (Figure 2E). Consistent with our findings, cuprizone treatment has been shown to increase oxidative stress, elevate production of reactive oxygen species (ROS), and induce endoplasmic reticulum stress, triggering UPR and cell death [24,25].

To validate our observations, we performed RNAscope of key genes represented in our enriched pathways in the DAO clusters. First, we tested IFN activation that was identified in LPC treatment using *Socs3* expression, a class I major histocompatibility complex (MHC-I) gene induced by cytokine signaling that, when knocked out, enhances MOL survival [26]. We observed a significant increase in the number of *Socs3*-expressing MOLs in LPC- and cuprizone-treated samples at remyelination, supporting our conclusion that both models converge on the MOL_E state at remyelination and suggesting these cells are less likely to survive and contribute to myelin repair (Figure 2G, H). To assess the specificity of the cuprizone-induced DAO state, we evaluated expression of *Cdkn1a,* a p53-dependent kinase that is induced in response to DNA damage that is highly specific to MOL_G. RNAscope confirmed *Cdkn1a* as a unique marker of cuprizone-induced OL stress (Figure 2I, J). Given that cuprizone treatment induces a more continuous and protracted OL loss, we wondered if a similar stress state is present at earlier stages of LPC-induced demyelination. However, analysis of LPC-treated samples at 2 days post lesion (dpl)

revealed no detected expression of *Cdkn1a* or *Nupr1* (another MOL_G marker), confirming that cuprizone-induced stress mechanisms are distinct from those induced by LPC (Figure 4K, L).

In addition to MOLs, we also examined OPC responses, given their critical role in remyelination. OPCs are known to migrate to sites of injury, induce immune activation, and participate in debris phagocytosis [27–29]. To characterize their response, we analyzed gene expression changes in OPCs in the LPC-induced demyelination time point, where progenitors are most enriched. We found that OPCs shared many DEGs and enriched pathways with both LPC- and cuprizone-treated DAO states (Supplementary Figure 4C, D). Immune-related and IFN response genes are upregulated, including *Ifit1*, *Stat2*, *Socs3*, and *C4b*. Additionally, genes associated with cell cycle arrest and DNA damage response, such as *Gadd45b* and *Gadd45g*, were also upregulated. We also observed downregulation of genes such as *Mertk*, which could indicate that OPCs are suppressing their phagocytic function in favor of differentiation. Together, these observations suggest that OPCs accumulate damage and promote inflammation, which could inhibit differentiation. Notably, *Socs3* expression is significantly upregulated at the remyelination time point in OPCs in both models, and *Cdkn1a* is upregulated in LPC-induced demyelination (Supplementary Figure 4E, F). These results support our hypothesis that OPCs share a similar state across models.

Overall, the activation of TNFα, IFN, Myc and p53 signaling pathways suggests that both models induce OL stress, albeit through distinct mechanisms. These stress-associated phenotypes resemble those previously described in MS, highlighting potential avenues for therapeutic intervention by targeting stress-response transcriptional changes in OPCs and MOLs [30]. Our results support the use of LPC and cuprizone models to study distinct aspects of OL stress and repair in CNS demyelination, and to identify strategies for enhancing OL protection, survival and remyelination.

Altered transcriptional states of oligodendrocytes are observed in MS lesions.

The observation of distinct DAO states in the mouse models, prompted us to next investigate whether these states resemble those found in MS lesions. To address this, we first subclustered and reannotated the human OL lineage cells (Figure 3A, Supplementary Figure 5A). Using known marker gene expression, we identified OPCs, COPs, and 4 MOL populations (MOL_A-G) (see Methods, Figure 3B, Supplementary Figure 5B). MOL_A is characterized by high *RBFOX1* and

SLC5A11 expression, and MOL_B additionally expresses low levels of *CTNND2* and *PLXDC2*. MOL_C loses expression of MOL_A markers, further upregulates MOL_B markers, and expresses *OPALIN* and *LAMA2*. MOL_D is distinguished from MOL_C by increased expression of *FTL* and *CSPRP1*. Additionally, we identified clusters that express markers for DAOs (MOL_E-G). Specifically, MOL_E expresses high levels of *HSPH1*, *FOS*, and *JUN*, while MOL_F and MOL_G are marked by *CDKN1A*, *NEDD4L*, and *GADD45B* expression. We further validated the DAO state using a gene signature from a previous MS patient study, confirming that known DAO genes are enriched in MOL_E-F and OPC clusters (Supplementary Figure 5C) [8]. Connectivity analysis revealed a primary differentiation trajectory from OPCs to COPs, followed by MOL_A through B, C, and D, with additional connectivity between the DAO states (Supplementary Figure 5D).

We hypothesized that DAO clusters would be enriched in MS lesions, rather than corresponding directly to patient diagnosis (Supplementary Figure 5E). Density plots revealed a broad shift in MOL_A-D states in MS lesions, but minimal differences between WM and NAWM samples (Figure 3C, Supplementary Figure 5F, G). As expected, OPCs are increased in ALs, and DAO clusters were enriched across the different lesion types. Examining the data by patient diagnosis, we found that DAO populations are primarily derived from secondary progressive Multiple Sclerosis (SPMS) and primary progressive Multiple Sclerosis (PPMS) patients, which make up the largest contribution of patient samples in the dataset (Supplementary Figure 5E). Interestingly, MOL_E- a comparatively small cluster- was highly enriched in two WM samples (44% of cells), both from 85-year-old individuals. MOL_E shares similar gene expression changes to the DAO, suggesting that normal brain aging may induce a stressed state in MOLs. Notably, we observed that MOL_A-D populations are more evenly distributed in human WM compared to the mouse corpus callosum, suggesting greater heterogeneity in adult human MOLs. Quantification using Milo supported our observations, showing minimal changes in homeostatic MOL_A-D in NAWM, but variable alterations in MS lesions, with the fewest changes observed in RLs (Figure 3D, Supplementary Figure 5H). Overall, all MS lesion types were strongly enriched for DAO states MOL_F and G.

To compare these DAO states to our mouse model findings, we performed differential gene expression analysis focused on AL and RL samples, as they are most comparable to our mouse time course data (Supplementary Data 4). Strong overlap in DEGs between AL and RL suggest that RLs retain transcriptional changes induced by demyelination rather than fully reverting to a healthy WM state (Figure 3E, F). However, we observed more transcriptional changes within AL

compared to RL, indicating that DAO cluster MOL_G shifts toward a more homeostatic-like state during remyelination. Pathway enrichment analysis revealed upregulation of IFNγ signaling in both AL and RL, while AL also exhibited enrichment in TNFα signaling, complement activation, and apoptosis pathways (Figure 3G). Importantly, many of these pathways were previously observed in the original analysis of this human dataset, including TNFα and IFN upregulation, supporting the robustness of our analysis pipeline [13].

Both LPC and cuprizone demyelination models share similarities with DAO states in MS lesions.

To determine how human DAO states are recapitulated in the mouse models of demyelination, we focused on the demyelination time points, as these encompass the majority of DEGs in the OL lineage. We found that the LPC-induced DAO states (MOL_E, F) shared 30 genes with the human DAO (MOL_G) cluster, including *ANXA2*, *APOE*, *STAT1*/*3* and *SNX10* (Figure 3H). *SNX10,* a member of the sorting nexin protein family, plays many roles across tissues, including proliferation of glioblastoma stem cells and phagosome maturation in macrophages, though its function in OL pathology remains unexplored [31,32]. We observed *SNX10* enrichment in the DAO and OPC clusters of human ALs, and particularly strong enrichment in LPC-treated MOL_E at remyelination, but little to no expression in cuprizone-treated cells (Figure 3I).

In contrast, the cuprizone DAO (MOL_G) cluster shared 34 upregulated genes with the human DAO cluster in AL, including *FOS*, *CDKN1A* and *SLC7A5* (Figure 3J). Notably, *CD44*, a transmembrane glycoprotein involved in cell adhesion and signaling, was highly enriched in the AL and cuprizone DAO clusters, but not in LPC demyelination (Figure 3K). Interestingly, ectopic expression of *CD44* in myelinated OLs induces progressive demyelination, suggesting that aberrant *CD44* expression may contribute to MS pathogenesis [33]. Further support of a shared DAO phenotype comes from the expression patterns of *ANXA2* and *SLC7A5,* which are similarly enriched in both MS lesions and distinct mouse models (Supplementary Figure 6A, B). Interestingly, only 2 genes, *LMNA* and *S100A6,* were shared across all conditions tested. *LMNA,* a key component of the nuclear lamina, has been shown to be required for the stability of the myelinating oligodendrocyte transcriptome, suggesting that lamin upregulation could be a common molecular change observed in OL damage [34]. *S100A6,* a calcium binding protein, is most known for its role in promoting proliferation and differentiation, and more recently for roles in stress response [35].

We also analyzed the overlap of downregulated genes, and found 94 shared DEGs between MS and cuprizone DAOs, but only 23 shared between MS and LPC DAOs (Supplementary Figure 6C, D). Many downregulated genes specifically shared between cuprizone and AL are involved in OL homeostasis and function, including myelin structural component genes *MOG*, *PLLP*, and *MOBP,* myelin stability factors such as *ERBB3* and *VLDLR*, and lipid metabolic genes such as *LPAR1* and *ACSS2*. Together, downregulation of these genes suggests that DAOs in the cuprizone model and in MS patients may have a compromised ability to produce stable myelin.

Finally, we explored the expression pattern of genes previously validated by RNAscope in our mouse models (see Figure 2). *CDKN1A* and *SOCS3* were highly enriched in the MS DAO cluster MOL_G, especially in RLs, indicating persistent activation of pathological pathways in MS lesions (Supplementary Figure 6E). Additionally, *NUPR1,* which was highly enriched in mouse cuprizone MOL_G cells, was also upregulated in MS lesions. Intriguingly, *NUPR1* expression is most highly enriched in NAWM, and appears to be more broadly expressed, suggesting its expression could be an early marker for inflammation in MS tissue (Supplementary Figure 6F). Together, these findings indicate that both mouse demyelination models exhibit some disease-associated characteristics of MS pathology in OLs, though their overlaps are largely non-redundant between the LPC- and cuprizone-induced DAOs. Future studies will be essential to understand the functional consequences of DAOs, to determine whether DAO states persist in mouse models as they do in post-mortem MS tissue, and to identify strategies to restore healthy MOL heterogeneity and promote remyelination.

Microglia become highly activated in response to demyelination injury and persist at remyelination

We previously observed that microglia undergo a dramatic state shift following injury, so we next sought to determine whether LPC and cuprizone treatments elicit different microglia responses. To address this, we subclustered and reannotated the microglia and macrophages, identifying 7 subclusters (Mg_A-E), BAMs and a cycling cluster (Figure 4A, Supplementary Figure 7A). Using a previously identified innate immunity gene set, we calculated a gene score and found enrichment of reactive microglia, termed disease-associated microglia (DAMs), within the highly-related Mg_C, D and E clusters, BAMs and cycling cells (Figure 4A, Supplementary Figure 7B) [36,37]. Further confirming this activated state, these clusters were enriched for known DAM markers

Apoe, Cd63, Lyz2, and downregulated homeostatic genes *Tmem119* and *P2ry12* (Figure 4B, Supplementary Figure 7C). BAMs expressed lineage markers *C1qa* and *Trem2,* but lack microglia-specific genes such as *Tmem119* and *Cxcr1*. Notably, BAMs also shared expression of many DAM markers, and additionally enriched for immune signaling genes *Igqap1* and *Ms4a7*.

To further investigate the microglia dynamics, we analyzed the distribution of homeostatic and DAM populations in response to LPC or cuprizone treatment. Baseline conditions were enriched for homeostatic microglia clusters (Mg_A and Mg_B) (Figure 4C, Supplementary Figure 7D, E). Cuprizone controls showed an increased proportion of Mg_B cells, however these cells exhibited a lowered gene count, suggesting a potential artifact from sample preparation (data not shown). LPC-induced demyelination drives a strong shift toward DAM states, cycling cells and BAMs, which are specific to LPC treatment. Cuprizone-induced demyelination also enriched DAM states, but retained more homeostatic microglia and fewer cycling cells, suggesting that cuprizone elicits a weaker microglial response compared to LPC. At remyelination time points, DAM states persisted, but resembled baseline more closely than at peak demyelination. Milo analysis supports these observations, showing that homeostatic clusters are lost mores strongly in LPC-treated samples compared to cuprizone treatment, and that DAM states are strongly enriched at demyelination and persist into remyelination for both treatments (Figure 4D, Supplementary Figure 7F). Notably, the DAM cluster Mg_D and cycling microglia are specific to demyelination in both treatments, indicating that these temporal shifts are a shared response to injury in both demyelination paradigms.

Toxicity-induced demyelination models induce transcriptionally similar microglia states

Although we observe highly comparable shifts in microglia states across treatments, we wondered if there are molecular differences between the DAMs induced by LPC or cuprizone treatment. To address this, we test for gene expression changes by comparing the homeostatic state (Mg_A) with DAMs, or BAMs from LPC-induced demyelination (Supplementary Figure 8A, Supplementary Data 5). To simplify the comparison, we combined Mg_C, D and E clusters to represent DAMs because DEG testing between these states showed largely overlapping profiles (Supplementary Figure 8B). In agreement with the strong state shift we observed, we found a large number of DEGs per condition. Comparing the DEGs between LPC and cuprizone treatments, we found a substantial overlap of both up- and downregulated genes (Figure 4E). Additionally, GSEA

revealed strong enrichment of many pathways in common, with almost all pathways shared between at least 3 of the tested states (Figure 4F). These pathways included metabolic pathways (e.g. cholesterol homeostasis, glycolysis, and oxidative phosphorylation), immune response pathways (e.g. IFN response, inflammatory response, and fatty acid metabolism), as well as hypoxia and ROS pathways, suggesting high metabolic demand may lead to ROS accumulation, potentially exacerbated by oxygen deprivation in the tissue. A mouse model of Alzheimer's disease previously showed that sustained hypoxia induces cell cycle arrest in activated microglia, and a shift to anaerobic glycolysis similar to what we observe in response to demyelination [38].

The most striking distinction we observed between treatments was the stronger enrichment of DAMs in LPC compared to cuprizone. We wondered if this enrichment would reflect a change in the magnitude of gene expression changes, that is not captured by only looking at the DEG overlap. To test this, we performed a pairwise log fold-change (logFC) analysis and fit a linear model to genes that are significantly differentially expressed in at least one of the conditions being compared. At demyelination, DEGs were more strongly up- or downregulated in LPC than in cuprizone (slope = 0.712, r^2=0.564) (Figure 4G, top). In comparison, the changes were more similar at remyelination across treatments (slope = 0.912, r^2=0.727) (Figure 4G, bottom). When comparing demyelination and remyelination within each treatment, we found that LPC-induced changes were more pronounced (slope=0.828, r^2=0.801), while cuprizone exhibited minimal differences between the time points (slope=1.002, r^2=0.893) (Supplementary Figure 8C). The largest change in magnitude was observed between BAMs and DAMs in LPC demyelination (slope = 0.59, r^2=0.702), further supporting BAMs role as potent phagocytic responders in LPC lesions.

To visualize these expression changes, we plotted a heatmap of pseudobulked gene expression of genes shared and uniquely differentially expressed across conditions (Supplementary Figure 8D). Genes upregulated across all BAM and DAM conditions showed highest expression during demyelination, returning toward baseline at remyelination. To further illustrate this trend, we plotted UMAPs with the expression of *Sash1,* an intracellular adapter protein and known tumor suppressor, which was among the most highly upregulated genes in DAMs. Interestingly, knockdown of *Sash1* was shown to enhance functional recovery in spinal cord injury models, underscoring the complex roles microglia play in both enhancing and inhibiting recovery [39–41]. We also plotted the expression of *Cacnb2*, a voltage-dependent calcium channel that is strongly downregulated in demyelination (Supplementary Figure 8E). Blocking calcium signaling has been

shown to be sufficient to enhance microglia transition to inflammatory phenotypes, and has been shown to be strongly downregulated in demyelination, consistent with our observations [42].

In summary, our findings reveal LPC induces a stronger microglial response than cuprizone, characterized by higher BAM infiltration, greater microglial proliferation, and sustained activation throughout remyelination. In contrast, cuprizone-induced demyelination elicits a more gradual response over weeks, leading to weaker DAM enrichment and a near-complete return to homeostasis by remyelination. Despite these differences, both models exhibit a highly conserved and persistent microglial response to demyelination, reinforcing their utility in studying microglial dynamics in MS pathology. However, our results highlight LPC as a model for studying robust microglial activation, including the recruitment of BAMs and stronger transcriptional shifts.

Human microglia exhibit higher heterogeneity in MS lesions than we observed in mouse demyelination models

To compare the microglia response in mouse models with MS, we characterized microglia states across MS lesions. Reclustering human microglia revealed 7 subclusters which we labeled Mg_A-G (Figure 5A, Supplementary Figure 9A). Scoring of pro-inflammatory gene sets revealed no enrichment in Mg_A, low enrichment in Mg_E, and strong enrichment in Mg_F, suggesting heterogenous DAM states in MS (Supplementary Figure 9B). Consistently, marker gene expression shows high expression of microglia lineage genes and low levels of inflammatory markers in Mg_A, which we consider homeostatic microglia (Figure 5B, Supplementary Figure 9C). In contrast, expression of genes involved in activated states are evident in the remaining microglia clusters, including *MYO1E*, *ALCAM*, *APOE*, and *HIF1A*. Mg_D strongly expresses genes previously used to describe 'foamy' microglia, such as lipid storage genes *LPL* and *PPARG,* and the lysosome gene *CTSL,* and supports a role for these cells in clearance of myelin debris [12]. Mg_E expresses similar gene expression to Mg_D, but also expresses *NUPR1*, a gene that is more abundant in white than grey matter, and is critical for repressing iron-dependent cell death [43,44]. Mg_F is more distinct from Mg_C and D, expressing the highest levels of *VEGFA* and *FOSL2*. Finally, Mg_G are peripheral macrophages, with high and specific expression of *F13A1*, *MRC1*, *CR1*, and *LYVE1*. Together, these findings suggest that MS induces greater cellular heterogeneity in the activation states than we observed in mouse demyelination models.

We next investigate whether MS lesions differentially enrich for activated microglia states. Density plots show that WM is composed almost entirely of homeostatic microglia (Mg_A). However, we observed state shifts from homeostasis were clearly evident in the NAWM, consistent with previous reports showing reactive microglia throughout the CNS in MS patients (Figure 5C, Supplementary Figure 9D, E) [43]. MS lesion types, however, show specific patterns of microglia activation. Quantification by Milo revealed a loss of homeostatic microglia across lesion types, except for RL, which shows almost no changes compared to WM (Figure 5D, Supplementary Figure 9F). AL and CAL samples are strongly enriched for Mg_D and Mg_E, suggesting foamy microglia are induced where we expect to observe the highest amounts of myelin debris. In contrast, CIL and particularly RL samples are enriched strongly for Mg_F, and largely lose Mg_D and E enrichment. This raises the possibility that remyelination-specific microglia may create an environment conducive to repair. Mg_F expresses *VEGFA*, which could promote angiogenesis in MS, and has been shown in mouse models to enhance OPC survival and differentiation [45,46]. We observed no changes in the macrophage cluster, Mg_G, which comprised very few cells in the analysis.

To define the transcriptional differences in activated microglia across distinct lesions, we calculated DEGs in Mg_D and Mg_E from ALs, and Mg_F from RLs by comparing the clusters to the homeostatic Mg_A in WM (Supplementary Data 6). We found several hundred genes differentially expressed in each condition. Among them, 176 upregulated and 133 downregulated genes were shared across all clusters, though many genes were specific to one state (Supplementary Figure 10A). GSEA pathway analysis revealed that TNFα signaling, hypoxia, and epithelial mesenchymal transition (EMT) are shared between Mg_D and F (Supplementary Figure 10B). Only EMT is enriched in Mg_E, which have high mitochondrial content and lower gene counts per cell, further suggesting they could be stressed or lower quality nuclei. Mg_D enriches additional pathways, including those related to immune functions, (e.g. inflammatory response and IFNγ signaling), and metabolism pathways (e.g. glycolysis and cholesterol homeostasis). In contrast, the remyelination-enriched Mg_F is enriched for coagulation and P53 signaling. We also examined whether these gene expression differences reflect patient-specific responses, as previously described [13]. To address this, we generated a heatmap of pseudobulk expression focused on DEGs in Mg_D, E, and F. Some variability was observed, particularly in RL, where the pseudobulk profile was based on only five samples, suggesting stratification of patient responses (Supplementary Figure 10C). Finally, we plotted the expression of several top DEGs, showing specificity in gene upregulation in the expected microglia states and further supports the

possibility these microglia exhibit functional differences, though this remains to be tested (Supplementary Figure 10D). Overall, these findings highlight greater cellular heterogeneity in human patient data compared to mouse models, suggesting that MS involves human-specific mechanisms not captured in demyelination models.

Human and mouse microglia show some conservation of activated transcriptional states in response to demyelination

Next, we compared human and mouse microglia states at the transcriptional level, focusing on the DAMs at the demyelination timepoints, where the strongest transcriptional response was observed. We found that mouse DAMs share a common but limited overlap of upregulated genes with the human Mg_D-F (Figure 5E). This overlap also does not appear to be more enriched in a treatment-specific manner, consistent with the conserved response observed when comparing the two treatment paradigms. Shared genes include 20 for LPC and 28 for cuprizone, with 18 that are common to both mouse treatments. This suggests a common core subset of genes that are highly conserved in activated microglia across species. However, we do not observe a preferential overlap of DEGs between mouse DAMs and a specific MS-enriched DAM cluster. These results suggest that mouse microglia exhibit less heterogeneity than those in human disease. While mouse microglia may perform multiple functions required to clear myelin debris and create a permissive environment for remyelination, these functions may be diversified in humans and restricted to specific subtypes, consistent with previous reports of evolutionary divergence in human microglia [47].

We examined the expression of some of the shared DEGs. Strong enrichment of the transmembrane protein *GPNMB* is observed, particularly in the mouse DAM state Mg_C and BAMs, and human Mg_E (Figure 5F, Supplementary Figure 10E). *COLEC12* is a cell surface protein and scavenger receptor, and has been shown to regulate microglia phagocytic activity, especially in myelin uptake [48]. We found that *COLEC12* is similarly enriched most strongly in mouse Mg_C and human Mg_E, but its expression was restricted to mouse microglia and not observed in the BAM cluster, though it is expressed in the human macrophage cluster (Figure 5G, Supplementary Figure 10F). *IQGAP1,* a scaffolding protein that has numerous roles in signaling, is more broadly enriched in human and mouse microglia and macrophages (Figure 5H, Supplementary Figure 10G) [49]. In summary, while there is some overlap in the differentially expressed genes and pathway enrichment between human and mouse microglia, the human

microglia exhibit more distinct states, suggesting higher heterogeneity in the DAM states compared to the mice, consistent with previous reports across species [47].

Cross-species integration reveals MS lesions and cuprizone treatment induce similar oligodendrocyte states.

Owing to the high cellular conservation across the mouse and human white matter demyelination samples, we sought to also characterize the cellular diversity by integrating the data. Using the sysVI method to correct for species-specific effects, we successfully integrated the total dataset, including both mouse and human samples (Supplementary Figure 11A) [50]. We observed species mixing across both cell types and single cell/nuclei, showing that integration achieved suitable batch correction while conserving biological variability. We confirmed that our annotation of cell types in our new integration matched the single species datasets using marker gene expression (Supplementary Figure 11B, C, also see Figure 1). We found that the BAM-specific expression of *F13A1* is integrated within the microglia cluster that expresses *CSF1R*, suggesting we lost the specification of macrophages and microglia that we observed in the mouse and human datasets. We examined the cellular states across lesion conditions and observed a modest transition in microglia and astrocytes, similar to the human-only integration (Supplementary Figure 11D). The previously described changes in glial cell abundance in each species are evident in our cross-species integration and correspond to expected changes based on lesion type (Supplementary Figure 11E). These quality metrics support our ability to use this integration to interrogate the similarities between MS and toxicity mouse models.

We next subset glial cells by their lineage and reclustered them. We obtained a high-quality clustering of OL lineage cells, but microglia did not separate into clusters based on our previously defined markers. This observation further supports our hypothesis that OLs have higher conservation across species than microglia. Thus, we focused on the OL cross-species comparison. We tested the replicability of cell types across our refined mouse and human OL clusters using MetaNeighbor analysis and found high similarity (higher AUROC score) between mouse and human cell types (Figure 6A). Specifically, mouse and human progenitors (OPCs, COPs, NFOLs and MFOLs) shared high similarity, and we observed broad clustering of two classes of MOLs: *OPALIN*$^{+}$ (human MOL_C/D and mouse MOL_B clusters) and *RBFOX1*$^{+}$ (human MOL_A and mouse MOL_C/D clusters). Human MOL_B and mouse MOL_A shared intermediate characteristics of these two MOL states, but appeared most similar to the *OPALIN*$^{+}$

cells. An additional grouping formed between human and mouse DAO states, supporting the shared similarities we described previously. Although cuprizone DAO state MOL_G clustered distinctly from the main DAO group, there was still high similarity with human DAO states (MOL_F and MOL_G clusters). Based on this replicability between species and marker gene expression, we annotated the reclustered cross-species OL UMAP by separating progenitors into OPCs and COPs, defining two broad classes of MOLs, one class of DAO, and one small cluster enriched in the MOL_E human WM-specific state (Figure 6B). We defined MOL_1 clusters with the highest expression of *RBFOX1*, *SLC5A11*, *RASGRF1, S100β,* and MOL_2 with *OPALIN*, *LAMA2*, *CTNND2, PLXDC2* (Figure 6C).

We focused on comparing the species contributions and transcriptional state of the DAO cluster, and found that this cross-species DAO state was most strongly enriched in cuprizone demyelination and MS lesions, especially AL (Figure 6D, E). Interestingly, we did not find the cross-species DAO cluster enriched in LPC treatment at either time point, but found that LPC-induced remyelination is enriched for cells that correspond to MFOLs and the MOL_E DAO state. Based on the LPC enrichment, we suggest that these findings support the previously proposed model that *OPALIN*$^{+}$ MOLs are in an intermediate state between progenitors and fully mature myelinating OLs [10]. Together, these observations suggest that cuprizone treatment induces a cellular state that shares greater similarity with the human DAO phenotype induced by active demyelination in MS. To further test this hypothesis, we created gene sets from the top DEGs specific to cuprizone and LPC demyelination DAOs (MOL_G and MOL_E, respectively). We confirmed enrichment of these gene sets in the different mouse treatments and tested the enrichment of these genes in the human-only embedding and mouse-human cross-species embedding (Figure 6F). In the human-only integration, we observed diffuse, low scoring cells for the LPC genes and stronger cuprizone scoring in the clusters corresponding to the human DAO states. The cross-species integration shows a similar pattern to the density plot enrichment, with weak LPC gene enrichment in the embedding corresponding to the *OPALIN*$^{+}$ MOLs, and stronger enrichment of the cuprizone-specific gene set in the shared DAO cluster. We plotted the expression of some of the treatment specific markers in the cross-species UMAP, supporting the cluster annotation and enrichment in the distinct demyelination conditions (Figure 6G). Notably, we also observe strong enrichment of both gene signatures in human OPCs, similar to the enrichment we see in the mouse models, suggesting there may be conserved mechanisms in the response of progenitors across species. We conclude that cuprizone treatment induces a state change in MOLs that more closely resembles active MS lesions than LPC.

Discussion:

The use of animal models to study demyelination and remyelination has been critical in advancing our understanding of OL susceptibility in neurodegenerative disease, and has been a valuable tool to address the urgent, unmet need for therapies promoting myelin protection and repair. While the neuroinflammatory EAE model is often considered the gold standard for MS, remyelination is difficult to analyze within this model, as lesions form sporadically throughout the spinal cord. In contrast, toxicity models, including cuprizone and LPC treatment, offer precise temporal control over damage and a predictable timeline for repair, making them particularly well suited for assessing therapies for halting disease progression and promoting remyelination. Numerous preclinical therapies have been identified using these models, with 25 out of 88 reported to have entered clinical trials and/or received approval [51]. However, whether distinct cellular responses are induced by different demyelination treatments remains unexplored, and these models have not been systematically compared to human patient samples across multiple cell types. Therefore, we analyzed single-cell/nuclei datasets to profile transcriptional changes in LPC- and cuprizone-induced demyelination and remyelination, comparing them to the largest MS dataset to date.

We found that LPC and cuprizone demyelination induce distinct OL phenotypes at the stages of demyelination studied. The only shared enrichment observed was in the TNFα signaling pathway. Blocking soluble TNF improved outcomes in EAE, suggesting that elevated TNFα signaling may have detrimental effects in both LPC and cuprizone models [52]. Importantly, we found that each model triggered unique transcriptional responses: LPC activated immune related genes involved in IFN signaling, whereas cuprizone induced a strong cellular stress response. Interestingly, Pandey et al. identified similar DAO states in their analysis of demyelination and Alzheimer's disease models, though they did not describe the dramatic and specific state change we observed in the cuprizone-induced MOL population [8]. These findings suggest that LPC and cuprizone models should be used strategically to investigate pathways enriched in each respective treatment. For example, cuprizone is the most widely used toxicity model for testing therapeutic interventions, yet whether drug targets have been more successful when aligned with the phenotypes we identified—such as P53 signaling, oxidative phosphorylation, or the unfolded protein response—remains an open question and a compelling avenue for future research [51,53].

Notably, LPC induced the most robust progenitor response, with OPCs exhibiting significant transcriptional differences. OPCs not only differentiate to generate new myelin but also play immune-modulatory roles in MS and EAE [10,28,54]. Although EAE is induced through immune modulation and is often considered more analogous to MS, we observed upregulation of multiple immune-related genes in MOLs within the toxicity models as well. For example, genes such as *Klk8*, *S100a10, H2-D1, Serpina3n, B2m* and other MHC-I genes were particularly upregulated in LPC treatment, where IFN signaling was also strongly induced. These findings indicate that OPC and OL immune responses are not exclusive to EAE-mediated demyelination.

Importantly, both LPC and cuprizone treatments recapitulate aspects of the human MOL phenotype observed in MS lesions. Cuprizone and AL DAOs exhibit shared upregulation of stress response genes, including *CDKN1A, NUPR1,* and *FOS*. Notably, MOLs in both cuprizone and MS lesions downregulate genes essential for myelin structure, stability, and lipid metabolism, reinforcing the idea that DAOs in MS exhibit compromised myelin production. This may explain why cross-species analyses reveal a stronger similarity between cuprizone-induced and MS DAO states.

Despite the distinct OL phenotypes observed during demyelination, both the cuprizone and LPC models converge on an altered differentiation state at remyelination stages, marked by the expression of stress response and immune-related genes, such as *Socs3*. The functional consequences of this MOL state at remyelination remain unclear. For example, it is not known if these OLs produce functional myelin that matures over time. A recent study demonstrated that even incomplete remyelination following cuprizone-induced demyelination is sufficient to restore neuronal function [2]. However, whether restoring proper MOL heterogeneity is essential for functional tissue regeneration remains an open question, requiring further investigation into how transcriptional phenotypes translate to MOL structure and function.

A critical factor in demyelination resolution is the role of microglia and macrophages. In addition to their homeostatic functions in the healthy brain, microglia respond to molecular damage signals and are essential for remyelination [55–57]. Both LPC and cuprizone treatments induce strong microglial activation phenotypes, with overlapping transcriptional changes affecting hundreds of genes. The striking number of shared genes across treatments and timepoints suggests a highly conserved demyelination response. LPC-mediated demyelination induces the strongest increase in cycling microglia and also leads to the accumulation of BAMs, potentially due to blood-brain

barrier damage from the injection, though no infiltration is observed in saline controls. Compared to cuprizone, LPC also sustains microglial activation longer during remyelination, whereas cuprizone treatment results in fewer DAMs and a greater proportion of homeostatic microglia. While no current therapies directly target microglia, some drugs with off-target microglial effects show regenerative benefits, suggesting that targeting microglia directly could enhance remyelination [57]. Our analysis supports the use of either toxicity model to study immunomodulatory effects on remyelination, though these phenotypes are not as heterogeneous as we observed in the MS dataset.

Microglia and macrophages in the MS cohort show important differences from the mouse models, particularly in their increased heterogeneity. Unlike in mouse models, human microglia do not show evidence of proliferation, aligning with previous findings that microglial populations are largely stable, with an estimated lifespan of ~4 years [58]. We identify a transcriptional signature similar to previously described foamy microglia (Mg_D), which are highly enriched in AL and CAL lesions. In contrast, microglia in remyelinating regions are largely homogeneous for the Mg_F state, which is characterized by strong enrichment for hypoxia and TNFα signaling. While some differentially expressed genes and pathways are conserved between human and mouse microglia, there is no clear enrichment for a specific human microglial subtype. Instead, the shared response primarily involves canonical demyelination-activated genes, such as *LPL, GPNMB,* and *APOE*. One potential explanation for the more homogeneous microglial response in mice is the shorter remyelination timeline (~3 weeks), compared to the prolonged disease course in MS. In longer-term neurodegenerative models, such as AD, distinct temporal microglial activation states emerge [59]. Comparing microglial responses in chronic or repeated demyelination models could reveal heterogeneity not captured in short-term studies and better align with the diversity observed in MS.

Although astrocytes are known to play a critical role in both demyelination models, insufficient astrocyte numbers were obtained for robust downstream analysis [17,18]. Additional enrichment steps are likely required to obtain astrocytes in sufficient numbers for downstream analysis in the mouse demyelination models [60] (Supplementary Figure 1F). In contrast, astrocytes constituted a significant proportion of the MS lesion samples (Figure 1H), leading us to subset astrocytes from the human WM dataset for subclustering and further annotation (Supplementary Figure 12). Strikingly, we observed astrocyte heterogeneity across MS lesions, similar to the changes we found in microglia, including a population of ciliated astrocytes that have been previously

described [15]. Future studies comparing astrocyte states between mouse models and MS lesions are necessary to fully understand the signaling mechanisms driving remyelination. For example, astrocytes have been shown to coordinate debris clearance via crosstalk with microglia [61]. Additionally, the blood-brain barrier plays a crucial role in damage signaling, and our analysis identified strong enrichment of innate immunity genes in endothelial cells and pericytes. Future studies investigating how cell-cell interactions regulate remyelination at the tissue level will be essential for understanding the resolution of demyelination and the mechanisms driving effective myelin repair.

Author information:

E.L.A performed the analysis of the single cell/nuclei RNAseq datasets, generated the figures, and wrote the manuscript. V.R.V and M.M.H performed the RNAscope validation experiments. K.L.A performed the single nuclei RNAseq experiments, conceptualized and supervised the project, and wrote the manuscript.

Acknowledgements:

We would like to thank Dr. Vittorio Gallo at Seattle Children's Hospital for his mentorship of K.L.A during early stages of this project and Dr. Cody Smith at the University of Notre Dame for critically reading this manuscript. We thank the Freimann Life Science Center at the University of Notre Dame for mouse breeding and husbandry, and Dr. Sara Cole from the Notre Dame Integrated Imaging Facility for her microscopy support. We thank Karuna Panchapakesan, Dr. Susan Knoblach, and Dr. Surajit Bhattacharya at Children's National Hospital for their assistance with single nuclei RNAseq library generation and analysis. Finally, we thank members of the Adams lab for their discussion. This work was supported by a career transition award to K.L.A. from the National Multiple Sclerosis Society.

Competing interests:

The authors declare no competing interests.

ORCID:

Erin L Aboelnour: https://orcid.org/0000-0002-1730-9405
Veronica R Vanoverbeke: https://orcid.org/0009-0008-2327-1344

Katrina L Adams: https://orcid.org/0000-0001-5290-4245

Data availability:

Mouse single nuclei raw data: *to be deposited in GEO.*
Pandey et al: Available on GEO (GSE182846).
Shen et al: Available on GEO (GSE148676).
Human MS dataset: Available on Zenodo (DOI: 10.5281/zenodo.8338963)

Code availability:

Before publication, representative python markdown and html scripts are being deposited in GitHub. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Figures

Figure 1: Conserved glial response in mouse and human white matter demyelination.

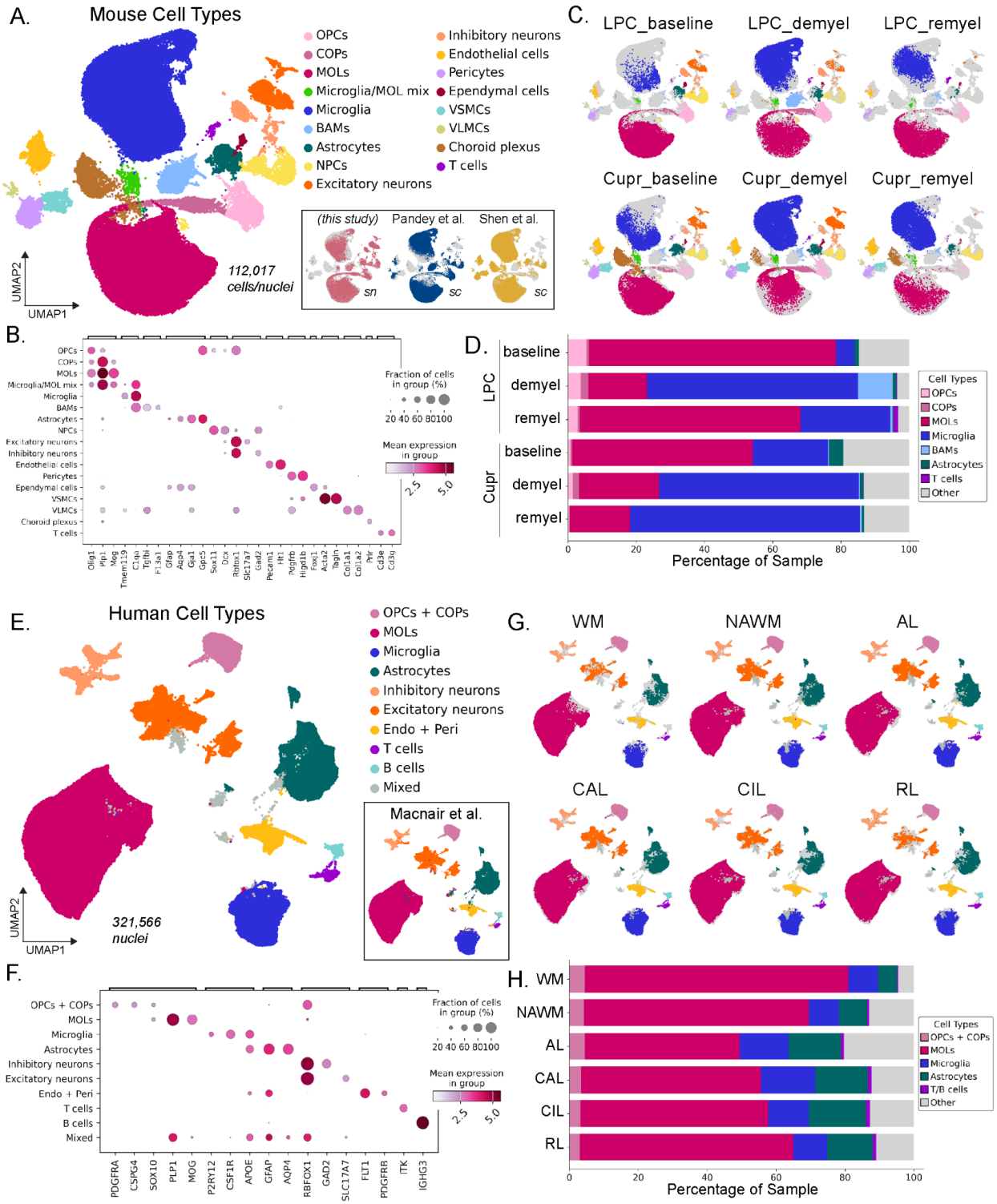

Figure 1: Conserved glial response in mouse and human white matter demyelination.

A. UMAP plot of single-cell/nuclei transcriptomic data from the microdissected corpus callosum of mouse demyelination models (Pandey et al. 2022; Shen et al. 2021). The inset UMAPs, colored by publication origin, demonstrate that sysVI integration produces a clear delineation of cell types and a high degree of overlap between sc/sn samples (Hrovatin et al. 2023).
B. Dot plot showing that the expression of selected cell type markers is restricted to the expected cluster annotations.
C. UMAP plots separated by demyelination treatment and time point, illustrating changes in glial cell abundance.
D. Stacked bar chart depicting the proportion of glial cell types per condition. (Other = NPCs, neurons, endothelial cells, pericytes, ependymal cells, VSMCs, VLMCs, and choroid plexus).
E. UMAP plot of human white matter control and MS samples. The inset UMAP demonstrates agreement between the annotated cell types and the original publication's annotations (Macnair et al. 2025). (Endo + Peri = endothelial cells and pericytes).
F. Dot plot showing the expression of selected cell type markers across annotated cell types.
G. UMAP plots separated by MS lesion type, revealing modest changes in glial cell states and abundances. (WM = white matter, NAWM = normal-appearing white matter, AL = active lesion, CAL = chronic active lesion, CIL = chronic inactive lesion, RL = remyelinated lesion).
H. Stacked bar chart showing the proportion of glial cell types per lesion type. (Other = neurons, Endo + Peri, and mixed cell types).

UMAP: uniform manifold approximation and projection, OPCs: oligodendrocyte progenitor cells, COPs: committed oligodendrocyte progenitors, MOLs: mature oligodendrocytes, BAMs: border-associated macrophages, NPCs: neural precursor cells, VSMCs: vascular smooth muscle cells, VLMCs: vascular and leptomeningeal cells, sn: single nuclei, sc: single cell, LPC: lysophosphatidycholine, Cupr: cuprizone, demyel: demyelination, remyel: remyelination, Endo + Peri: endothelial cells and pericytes, WM: white matter, NAWM: normal appearing white matter, AL: active lesion, CAL: chronic active lesion, CIL: chronic inactive lesion, RL: remyelinated lesion

Figure 2: LPC and cuprizone treatment induce distinct transcriptional states in oligodendrocytes

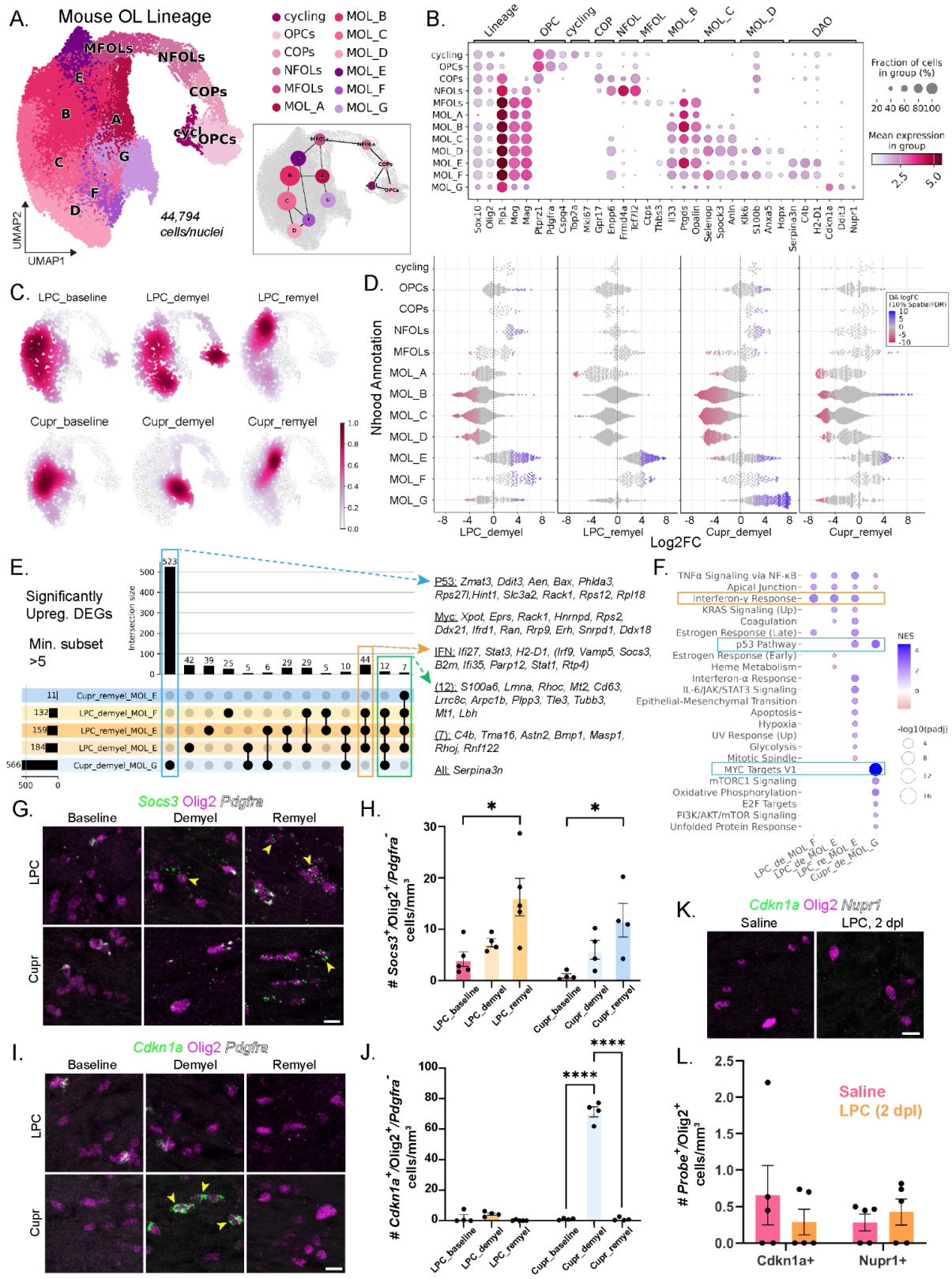

Figure 2: LPC and cuprizone treatment induce distinct transcriptional states in oligodendrocytes.

A. UMAP embedding of subset and reclustered mouse oligodendrocyte lineage cells. The inset UMAP shows PAGA connectivity with edge weights above 0.4 between the annotated cell types.
B. Dot plot displaying markers used to annotate the cell types.
C. UMAP plots colored by cell density in the embedding space, separated by treatment and time point. Data were downsampled to an equal number of cells per condition (n = 1,000 cells/nuclei).
D. Beeswarm plots showing differential abundance (DA) in neighborhoods (nhoods) assigned to the annotated cell types for each condition and time point. Comparisons were calculated against the corresponding baseline samples for each treatment using Milopy (Dann et al. 2022). Blue dots represent nhoods significantly increased in the sample, while magenta dots correspond to nhoods significantly decreased. (Spatial false discovery rate (FDR) < 0.1).
E. Upset plot illustrating the unique (single points) and overlapping (connected points) significantly upregulated genes (logFC ≥ 1.5, padj < 0.05) in the indicated condition and MOL subtype comparison (see text and Methods for description). Gene lists were subset to only include genes that were tested in all datasets, and overlap size was filtered to include only points with 5 or more genes. Colored boxes and corresponding dashed arrows highlight differentially expressed genes associated with the indicated pathways in the unique or overlapping gene set, highlighted in the GSEA plot (see Figure 2F).
F. Dot plot displaying Hallmark pathways enriched in the indicated condition and cell type, colored by the normalized enrichment score (NES) (see Methods for details). Colored boxes highlight gene sets corresponding to those in the upset plot (Figure 2E).
G. Representative RNAscope images of the mouse corpus callosum showing *Socs3* expression in MOLs. Yellow arrows point to *Socs3*$^{+}$ oligodendrocytes. (LPC: baseline = saline, demyel = 7 dpl, remyel = 28 dpl; Cuprizone: baseline = untreated, demyel = 5 weeks, remyel = 5 weeks + 3 weeks of recovery). (Scale bar = 10 µm).
H. Quantification of the total number of *Socs3*$^{+}$ oligodendrocytes per area across lesion conditions (n = 4-5 animals per group, one-way ANOVA with Tukey's multiple comparisons adjustment, LPC: *p = 0.0115, Cupr: *p = 0.0164).

I. Representative RNAscope images from the mouse corpus callosum showing *Cdkn1a* expression in MOLs. Yellow arrows point to *Cdkn1a*$^+$ oligodendrocytes. Time points are the same as in Figure 2H. (Scale bar = 10 μm).
J. Quantification of the total number of *Cdkn1a*$^+$ oligodendrocytes per area across lesion conditions (n = 4 animals per group, one-way ANOVA with Tukey's multiple comparisons adjustment, ****p-value < 0.0001).
K. Representative RNAscope images of the mouse corpus callosum showing lack of *Cdkn1a* and *Nupr1* expression in oligodendrocyte lineage cells in saline control (left) and 2 dpl after LPC injection (right, scale bar = 10 μm).
L. Bar chart quantifying the total number of *Cdkn1a*$^+$ and *Nupr1*$^+$ oligodendrocyte lineage cells at 2 dpl (n = 5 animals).

(NFOL: newly formed oligodendrocyte, MFOL: myelin-forming oligodendrocyte, DAO: disease-associated oligodendrocyte, Nhood: neighborhood, DA logFC: differential abundance log fold change, FDR: false-discovery rate, Upreg: upregulated, DEGs: differentially expressed genes, Min: minimum, resp: response, INF: interferon, NES: normalized enrichment score, padj: p-adjusted value)

Figure 3: Mouse demyelination models share transcriptional phenotypes with disease-associated oligodendrocytes in human MS Lesions.

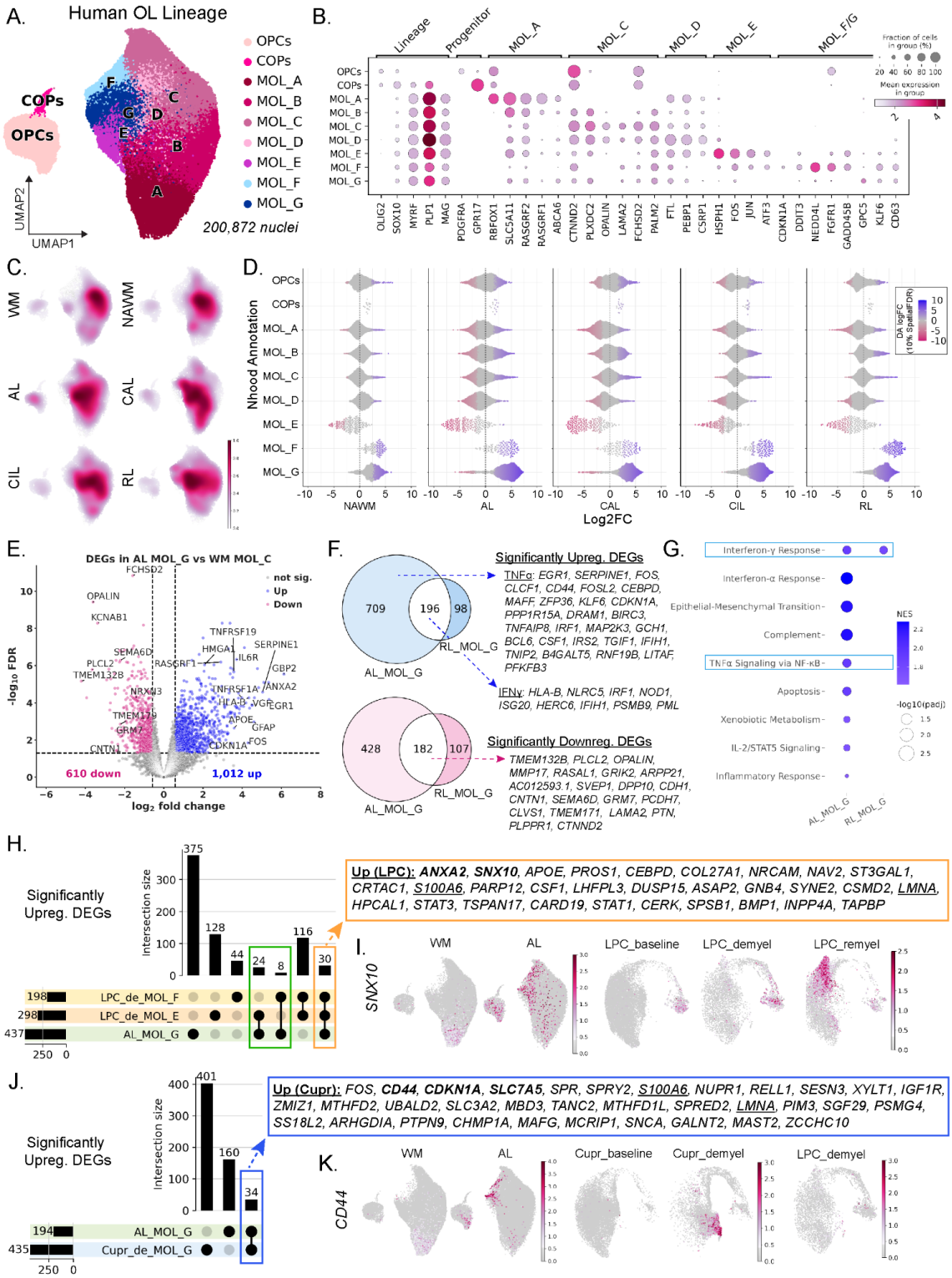

Figure 3: Mouse demyelination models share transcriptional phenotypes with disease-associated oligodendrocytes in human MS Lesions.

A. UMAP embedding of subset and reclustered human oligodendrocyte lineage cells.
B. Dot plot displaying cell type markers used to annotate the clusters.
C. UMAP plots colored by cell density in the embedding space, separated by lesion type. Data were downsampled to an equal number of cells per condition (n = 10,000 nuclei).
D. Beeswarm plots showing DA nhoods assigned to the annotated cell types for each lesion type, compared to WM using Milopy. Blue dots represent nhoods significantly increased in the sample, while magenta dots correspond to nhoods significantly decreased. (Spatial FDR < 0.1).
E. Volcano plot showing the statistical significance (y-axis) and logFC (x-axis) of all genes tested for differential expression in AL MOL_G compared to WM MOL_C. Significantly upregulated genes (logFC ≥ 1.5, padj < 0.05) are colored blue, and downregulated genes (logFC ≤ 1.5, padj < 0.05) are colored magenta, and labeled with the total number of DEGs. Dashed horizontal and vertical lines indicate the significance thresholds for padj and logFC, respectively.
F. Venn diagrams showing the overlap in significant DEGs in AL and RL, comparing MOL_G from the lesion and MOL_C in WM. A subset of genes are highlighted, corresponding to the given intersection by the arrow. Genes include are specifically upregulated in AL and involved in TNFα signaling, and genes that are up- and downregulated in common, including genes involved in IFN response.
G. Dot plot displaying Hallmark pathways enriched in the indicated condition and cell type, colored by the NES.
H. Upset plot illustrating the unique (single points) and overlapping (connected points) significantly upregulated genes (logFC ≥ 1.5, padj < 0.05) in the AL condition and LPC demyelinated MOL comparisons. Colored boxes highlight shared DEGs across species, with the upregulated genes shared by all conditions listed to the right. Bolded genes indicate genes that have been plotted to examine expression in lesion conditions, and underlined genes indicate genes shared in both human and mouse across all comparisons tested.
I. UMAP plots showing *SNX10* expression in WM and AL, with corresponding expression of the homologous gene in the LPC mouse model.

J. Upset plot illustrating the unique (single points) and overlapping (connected points) significantly upregulated genes (logFC ≥ 1.5, padj < 0.05) in the AL condition and cuprizone demyelinated MOL_G cluster. The blue colored box highlights shared DEGs across species, with the genes listed to the right. Bolded genes indicate genes that have been plotted to examine expression in lesion conditions, and underlined genes indicate genes shared in both human and mouse across all comparisons tested.

K. UMAP plots showing *CD44* expression in WM and AL, with corresponding expression of the homologous gene in the cuprizone mouse model, and LPC demyelination.

Figure 4: Microglial responses to demyelination are highly conserved across mouse models.

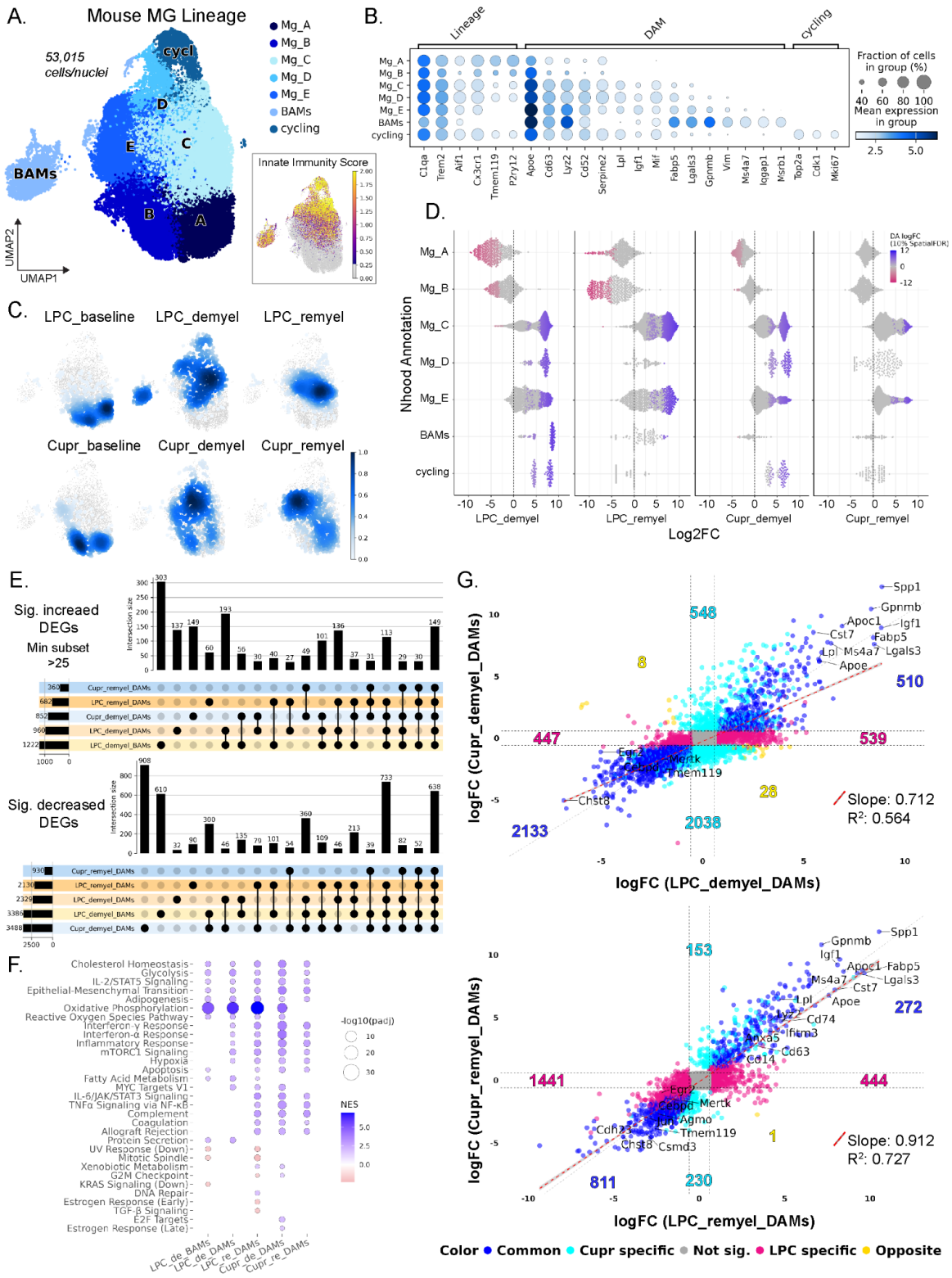

Figure 4: Microglial responses to demyelination are highly conserved across mouse models.

A. UMAP embedding of subset and reclustered mouse microglia and macrophage lineage cells. The inset UMAP plot shows a gene score calculated from the expression of genes involved in innate immunity (see Methods). (Mg = microglia, BAMs= border-associated macrophages).
B. Dot plot displaying cell type markers used to annotate the clusters. (DAM = disease-associated microglia).
C. UMAP plots colored by cell density in the embedding space, separated by treatment and time point. Data were downsampled to an equal number of cells per condition (n = 861 cells/nuclei).
D. Beeswarm plots showing DA nhoods assigned to the annotated cell types for each condition and time point. Comparisons were calculated against the corresponding baseline samples for each treatment using Milopy. Blue dots represent nhoods significantly increased in the sample, while magenta dots correspond to nhoods significantly decreased. (Spatial FDR < 0.1).
E. Upset plots showing the unique (single points) and overlapping (connected points) significantly upregulated (logFC ≥ 1.5, padj < 0.05, top) and downregulated (logFC ≤ 1.5, padj < 0.05, bottom) genes. Gene lists were subset to only include genes that were tested in all datasets. Differential expression was calculated for BAMs and DAMs (the combined pseudobulk of Mg_C, D, and E) and compared to Mg_A. Overlap size was filtered to include only intersections with at least 25 genes.
F. Dot plot displaying significant GSEA terms in the indicated condition and cell type, colored by the NES.
G. Scatter plots showing a 4-way comparison of the logFC of genes in LPC (x-axis) and cuprizone (y-axis) tested for differential expression in demyelination (top) or remyelination (bottom). Each dot represents one gene and is colored based on whether or not it overlaps with the comparative dataset. Grey dots represent genes that are not significantly differentially expressed in either condition. Blue indicates significantly up- or downregulated genes shared in both LPC and cuprizone. Cyan denotes genes that are significantly different only in cuprizone treatment, while magenta represents genes specific to LPC treatment. Yellow indicates genes significantly differentially expressed in both conditions but changing in opposite directions. Vertical and horizontal dashed black

lines represent the logFC cutoff of 1.5 fold. The grey dashed line represents x=y, and the red dashed line represents the linear trend line shaded with the confidence interval in grey, calculated between genes that are significantly differentially expressed in at least one condition. Each plot is labeled with the slope and R^2 of the linear trend line.

Figure 5: Human microglia exhibit distinct states across MS lesions and conserved features to mice.

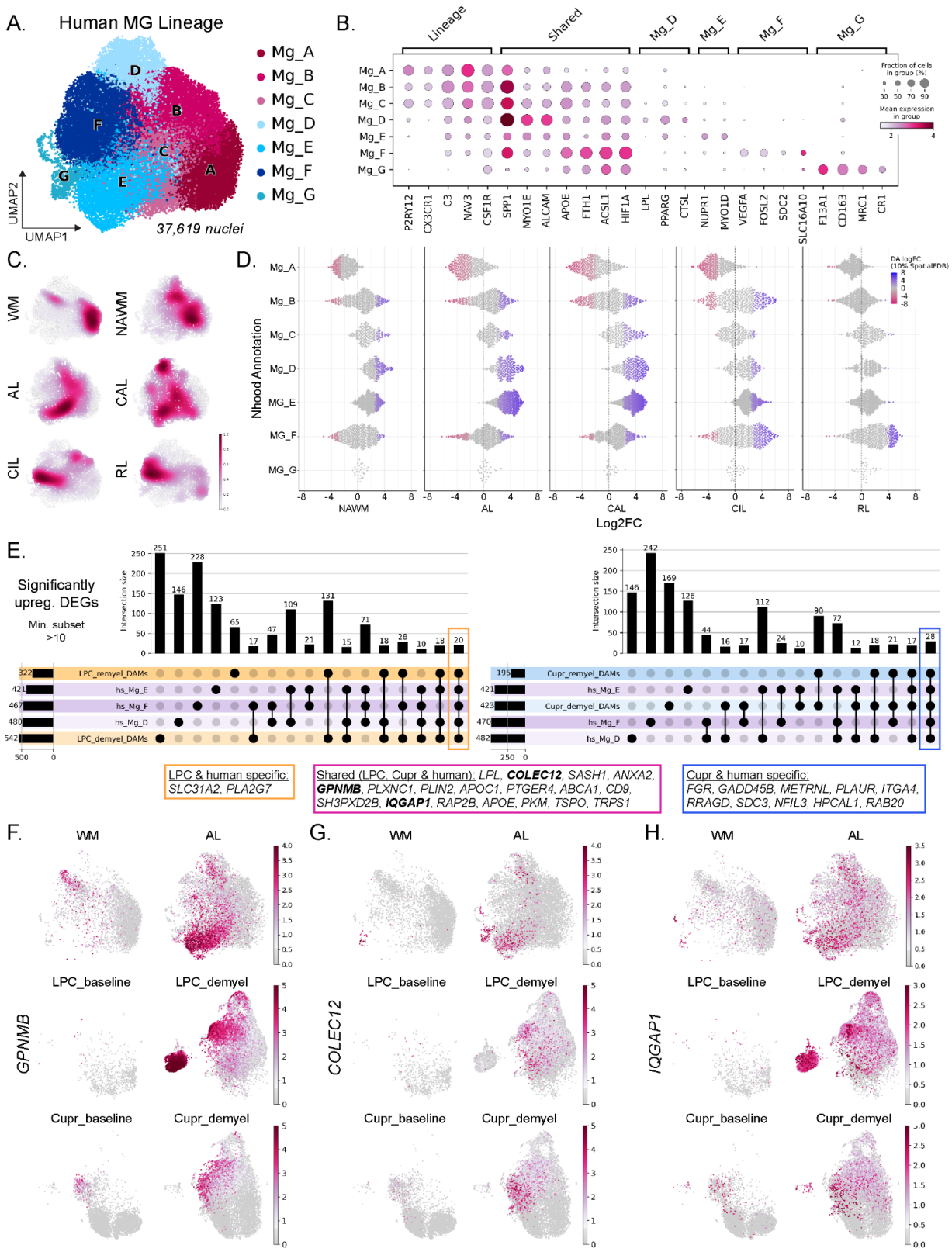

Figure 5: Human microglia exhibit distinct states across MS lesions and conserved features to mice.

A. UMAP embedding of subset and reclustered human microglia lineage cells.
B. Dot plot displaying cell type markers used to annotate the clusters.
C. UMAP plots colored by cell density in the embedding space, separated by lesion type. Data were downsampled to an equal number of cells per condition (n = 4,234 nuclei).
D. Beeswarm plots showing DA nhoods assigned to the annotated cell types for each lesion type, compared to WM using Milopy. Blue dots represent nhoods significantly increased in the sample, while magenta dots correspond to nhoods significantly decreased. (Spatial FDR < 0.1).
E. Upset plot illustrating the unique (single points) and overlapping (connected points) significantly upregulated genes (logFC ≥ 1.5, padj < 0.05) in the indicated MS lesion condition and cell type, and mouse DAMs for LPC treatment (left) and cuprizone treatment (right). Overlap size was filtered to include only intersecting points with 10 or more genes. Colored boxes highlight shared DEGs across all conditions, and are labeled below. Genes that are shared across both LPC and cuprizone with the human cells are labeled in the magenta box.
F. UMAP plots showing *GPNMB* in human WM and AL, and mouse controls and demyelination time points.
G. UMAP plots showing *COLEC12* in human WM and AL, and mouse controls and demyelination time points.
H. UMAP plots showing *IQGAP1* in human WM and AL, and mouse controls and demyelination time points.

Figure 6: Cross-species oligodendrocyte analysis reveals a conserved state between MS lesions and cuprizone demyelination.

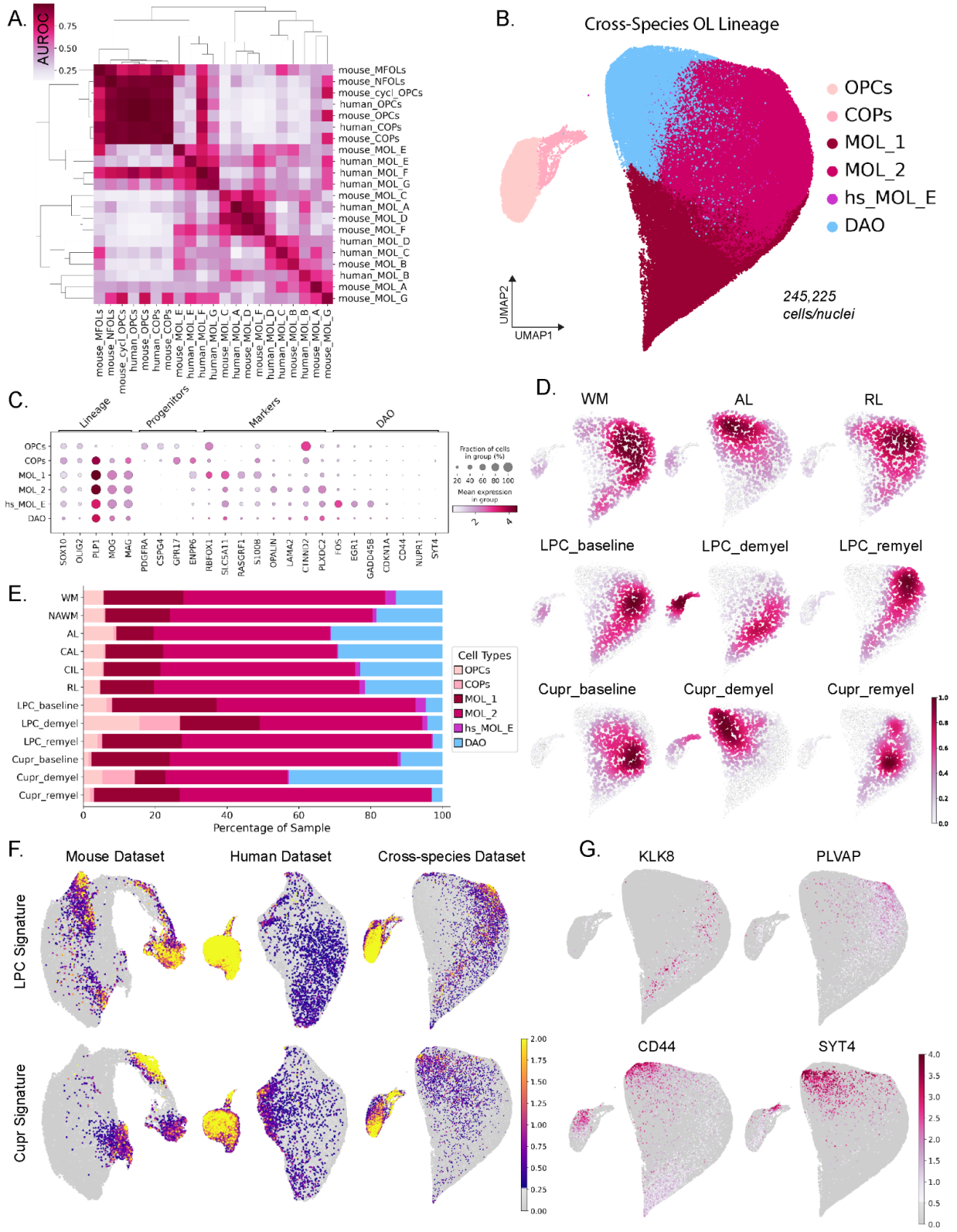

Figure 6: Cross-species oligodendrocyte analysis reveals a conserved state between MS lesions and cuprizone demyelination.

A. Heatmap of the mean area under the receiver operator characteristic curve (AUROC) scores with hierarchical clustering between human and mouse refined oligodendrocyte lineage cell types before cross-species integration. More similar cell types have a higher score.
B. UMAP embedding of subset and reclustered human and mouse oligodendrocyte lineage cells.
C. Dot plot displaying cell type markers used to annotate the cell types.
D. UMAP plots colored by cell density in the embedding space, separated by lesion type. Data were downsampled to an equal number of cells per condition (n = 1,000 nuclei).
E. Stacked bar chart depicting the proportion of cell types per condition.
F. Gene score of LPC and cuprizone enriched genes (see Methods) in mouse dataset (left), human dataset (middle) and cross-species integration (right).
G. Expression of genes specific to LPC (top) or cuprizone (bottom).

Methods

Experimental procedures:

Animals:

Wild-type C57BL/6 mice were maintained in the animal facility at the Children's National Hospital and the University of Notre Dame according to the Institutional Animal Care and Use Committee and the National Institutes of Health guidelines. Both male and female mice were analyzed for all experiments. Mice were maintained in the animal facility under a 12 h dark–light cycle, and constant temperature (20–26 °C) and humidity maintenance (40–60%). All studies complied with all relevant animal use guidelines and ethical regulations.

Demyelination protocols:

For LPC demyelination, adult (8-12wk) C57BL6 mice were anesthetized with ketamine/xylazine, the skull exposed, and a small hole drilled 1 mm rostral and 1 mm lateral of bregma. The animal was placed in a stereotactic rig and injected with 1uL of 1% LPC into the corpus callosum of one brain hemisphere at 2 mm depth over 5 minutes. Saline was injected on the contralateral side using the same steps as above. Mice were sacrificed as described below at 7 dpl for snRNA-seq and RNAscope for the demyelination timepoint, and at 28 dpl for RNAscrope at the remyelination timepoint.

For snRNA-seq, adult (8-12wk) cuprizone demyelination was conducted by feeding adult (8-12wk) C57BL6 mice 0.2% cuprizone chow for 5 weeks. For RNAscope validation, mice were fed 0.3% cuprizone chow for 5 consecutive weeks, and then either sacrificed as described below at the end of 5 weeks for the demyelination timepoints, or returned to normal chow for 3 weeks and collected for the remyelination timepoint. Control mice were fed normal chow and processed as described below.

Dissection and single nuclei preparation:

For LPC demyelination snRNA-seq collection at 7 dpl, mice were IP injected with neural red dye 2-3 hrs prior to tissue collection to label the lesion within the corpus callosum. Mice were then anesthetized, perfused with cold PBS for 5 mins, and the brain was dissected out into cold PBS. Coronal slices of the brain were made, to identify the lesion location, and the corpus callosum lesions were dissected out. For LPC controls, the corresponding contralateral side with saline injection was also collected. For cuprizone fed mice, the whole corpus callosum was dissected out. Tissue was placed in Eppendorf tubes, excess liquid was removed as much as possible, and tissue was flash frozen in liquid nitrogen. Frozen tissue was stored at -70°C.

Nuclei were pooled from 5 female and 5 male mice per condition for each replicate, with two replicates per condition, (LPC: baseline = saline injected, demyelination = 7dpl, Cuprizone: baseline = normal chow, demyelination = 5wks of 0.2% cuprizone). Pooled tissue was transferred to chilled glass dounce with 1mL of lysis buffer + 1% TritonX100 + 1 μM DTT (Sigma Nuclei Pure Prep Kit, Cat #NUC-201). Samples were dounced on ice 10 times with pestle A, then 10 times with pestle B. The suspension was passed through a 40μM FlowMi cell strainer. Dounces were rinsed with 1mL of low sucrose buffer and passed through the cell strainer as well. Samples were centrifuged at 300 rcf, at 4°C, for 5 mins in a fixed-angle centrifuge. Supernatant was carefully

decanted, and the pelleted nuclei were washed with 1.5mL of low sucrose buffer, and centrifuged again at 300 rcf, at 4°C, for 5 mins. Pelleted nuclei were then resuspended in 500µL Nuclei wash buffer (PBS + 1% UltraPure BSA + 0.2U/µL Protector RNase inhibitors) + 900µL of 1.8M sucrose cushion buffer, then carefully layered on top of 500µL 1.8M sucrose cushion buffer in a 2mL low-bind tube. Samples were centrifuged at 13,000 rcf, at 4°C, for 45 mins. Supernatant was removed, and nuclei were resuspended in 1mL of Nuclei wash buffer, passed through a 40µM FlowMi cell strainer and counted using a Cell Countess II. Approximately 6,000 nuclei were loaded per sample onto the 10X chromium controller per manufacturer's instructions. Single nuclei libraries were generated using the 10X single cell 3' reagents and protocol (v3.1). Libraries were sequenced on an Illumina Novaseq6000, with an average number of 80,000 reads per cell for each sample.

RNAscope and immunohistochemistry (IHC) validation:

Mice were anesthetized using isoflurane and intracardially perfused with cold PBS followed by 4% PFA for 5 minutes each. The brain was dissected out into cold 4% PFA and postfixed overnight. The tissue was then cryoprotected in 30% sucrose in PBS. Brains were embedded in OCT and frozen on dry ice. Sections were collected onto slides (12µM) and stored at −70°C.

IHC was performed first by thawing the slides (all steps are performed at RT unless otherwise indicated) in a humidifying tray for 5mins, permeabilized for 15 mins with 1xPBS + 0.1% tween (PBST), and blocked in blocking buffer (PBST + 1% normal donkey serum (NDS)) for 30 mins. Primary antibody (rabbit anti-Olig2, Millipore AB9610, 1:300 or goat anti-Olig2, R&D Systems AF2418, 1:100) was incubated in blocking buffer overnight at 4°C. Slides were washed 3 time for 5mins each with PBST, and incubated with the indicated fluorophore-conjugated secondary antibodies (AlexaFluor) in blocking buffer for 2 hrs. Tissue was washed again 3 times with PBST, and then post-fixed with 4% PFA for 10 mins. Next, the RNAscope Multiplex Fluorescent Reagent Kit V2 (Advanced Cell Diagnostics #323100) was used to perform fluorescent labeling of the indicated RNA molecules according to the kit manual, beginning with protease treatment. RNAscope probes used were: *Cdkn1a* (408551-C2), *Socs3* (437741), *Pdgfra* (480661-C3), and *Nupr1* (434811). Following RNAscope, sections were coverslipped with mounting media (Southern Biotech). Sections were imaged using the Leica Stellaris 8 DIVE confocal microscope and cell counting was analyzed for at least 3 sections per animal. Cells with at least two

fluorescent puncta were counted as positive for that probe. Olig2$^+$ cells positive for *Pdgfra* were counted as OPCs, and Olig2$^+$ cells negative for *Pdgfra* were counted as MOLs.

Data analysis:

Single cell and nuclei raw data processing:

Mouse: Demultiplexed fastq files were aligned to mm10 using Cell Ranger (v6.0.1), and published datasets were downloaded from GEO, and mapped using Cell Ranger (v7.2.0) [7,8]. Mapped h5 files were run through CellBender to remove ambient RNA [62]. Corrected count matrices for each sample were processed using scanpy tools and recommended best practices by retaining cells for downstream analysis [63]. Doublet detection and removal was performed using Solo, by first training a scVI model on each batch and then removing predicted doublets [64]. After doublet removal, data was merged across its publication of origin (ie. unpublished single nuclei data, Pandey et al, or Shen et al). Each of these 3 datasets was then subset based on quality metrics for that dataset- cells or nuclei were removed if they had less than 200 genes identified, more than 20% ribosomal content, or more than 5% mitochondrial reads. Additionally, we filtered the data to remove the highest outliers using the following criteria: the single nuclei data was filtered to remove samples with more than 15,000 total reads, or more than 7,000 genes. Pandey et al single cell data was subset to remove cells with more than 8,000 genes per cell, or 50,000 reads per cell. Shen et al was subset to remove cells with more than 6,000 genes identified, or more than 20,000 reads per cell.

Human: The MS data quality control counts matrix was downloaded from Zenodo (DOI: 10.5281/zenodo.8338963), subset to only the white matter samples, and converted to an h5ad file [65]. A major advantage of this dataset is that it contains carefully staged lesions. In brief, the subset data comprises 92 samples collected from 47 individuals, including 13 control (Ctr) and 34 MS patients, categorized by disease progression as relapse remittent (RR), secondary progressive (SP), or primary progressive (PP) MS (Supplementary Figure 2D). The WM only subset contains 13 controls (WM, mean age at death: 64.8 years, median: 70 years), 17 normal-appearing white matter samples (NAWM, mean: 61.3 years, median: 66 years), 21 active lesions (AL, mean: 51.5 years, median: 47 years), 17 chronic active lesions (CAL, mean: 54.7 years, median: 48 years), 13 chronic inactive lesions (CIL, mean: 65 years, median: 69 years), and 11 remyelinated lesions (RL, mean: 66.2 years, median: 69 years). Samples were further controlled

by removing nuclei with greater than 10% mitochondrial counts, more than 10,000 genes per cell, and more than 50,000 total reads.

Dataset integration and annotation:

Mouse: First, each dataset was integrated within its own experimental group (ie. each data source separately) to obtain harmonized cell type labels using scVI [66]. Briefly, we subset the data to the top 5,000 highly variable genes (HVGs), with the batch key set to each individual batch, and the percent counts for mitochondria and ribosome as continuous covariates. The latent representation from the trained scVI model was used to calculate the neighborhood graph, and then the UMAP embedding. Cell types were labeled by visual inspection of known marker gene expression, using markers calculated using scanpy's rank_genes_groups, and using custom cell type scoring using the marker gene sets detailed in Koupourtidou et al (see also lists under Gene Set Scoring). For full integration of the mouse single cell and nuclei data we used sysVI, which was developed to handle challenging integration tasks, such as between single cell and single nuclei data or between species [50]. For the mouse only integration we used the top 3000 HVGs, using the scanpy highly_variable_genes function set to 'cell_ranger', batch key set to the system variable (single cell or single nuclei), and each of the individual samples as a categorical covariate. We verified marker gene expression was enriched in the cell type labels, and that each cell type contained single cells and nuclei.

Human: To match our mouse corpus callosum datasets, we integrated the WM samples only using scVI, and produced a similar embedding and annotation as the published dataset [67]. For the mode, the batch key was set to the sample_id (representing each individual sample), and RNA count and percent mitochondrial count were given as categorical covariates. For training the model we used n_layers = 2, n_latent = 30, and gene_likelihood = 'nb'.

Cross-species: We also used sysVI to perform our cross-species analysis, integrating across all mouse and human cell types. First, we mapped the gene orthologs by downloading the stable gene ID and gene names from biomart using GRCh38.p14 and GRCm39. Shared orthologues were retained for a total of 13,867 genes in our dataset. We calculated the top 5,000 HVGs for each species, and then took the intersection, resulting in 2,703 shared genes. The sysVI batch key was set to a system variable designating single-nuclei human, single-nuclei mouse, and single-cell mouse. The sysVI categorical covariate key was set to each individual sample. The

quality of the integration was visually inspected to see overlap between datasets within each annotated cell type.The latent representation of the sysVI model was then used to calculate the neighborhood graph and embedding. Recent recommendations in atlas building can be found here [68].

Differential abundance analysis using Milopy:

Milopy, the python implementation of the miloR toolset, was used to calculate the differential abundance (DA) between neighboring cells [69]. Using the embedding calculated for the cell subtypes analyzed above (namely, oligodendrocyte and microglia lineages), the data was converted to a MuData object. For the mouse oligodendrocyte lineage, neighborhoods (nhoods) were recalculated using the X_sysVI representation and n_neighbors set to 100, for mouse microglia set to 300, for human oligodendrocytes set to 200, and for human microglia to 100. The make_nhoods function was applied with proportion set to 10% to create nhoods, and the median was inspected to ensure that there are on average at least 5 cells per sample per nhood. Pairwise comparisons were performed to calculate DA, and build an nhood_graph. The nhoods were annotated by the cell type labels by calculating the fraction of cells in a hood, if it was more than 60% a single cell type. For nhoods that were less than 60% a single cell type, the annotation was updated to 'Mixed', though this category is not shown in the final beeswarm plots. The final calculation and annotation was saved from the milo object, and used for input to plotting functions in ggplot2.

Differential expression analysis:

<u>Mouse:</u> Differential gene expression was calculated by pseudobulking, summing raw gene expression counts in each replicate, for each cell type in each replicate. Pseudobulked replicates include only aggregation of at least 15 cells. Pseudobulk counts were then subset to test individual pairwise comparisons, (ie. LPC baseline vs LPC demyelination). For each comparison, the counts matrices were trimmed using the edgeR filterByExpr function, using the design matrix as an input, and keep.lib.sizes set to FALSE. Each test was performed separately because not all samples contain single cell and nuclei samples, so batch correction was performed differently, but filtering was also done separately so the list of all tested genes is different between samples, as can be seen in supplementary files. Normalization factors were calculated using the TMM method. For mixed samples (with cells and nuclei) we used voom with quality weights and duplicate correlation

twice, blocked on the library type [70]. For samples with only cells, we used voom with quality weights and no blocking. For each treatment condition, we made pairwise comparisons of DAO clusters to their closest baseline MOL state (*MOL_E vs. MOL_B, MOL_F vs. MOL_D, and MOL_G vs. MOL_A*). Analysis was not performed on subclusters that had too few cells or only 2 replicates. In particular, we were unable to obtain differential testing in the cuprizone remyelination timepoint, due to low cell number (2,407 across 3 replicates), small library sizes, and high heterogeneity between replicates. Using lmFit, a linear model was fitted that included blocking on the batch key and correlation set to the duplicateCorrelation output. Empirical Bayes smoothing was applied, and the results table extracted using topTable with n=Inf to include all genes in the output. The full differential expression output for each cell type tested is included as a supplementary table. Unless otherwise noted in the figure legend, genes are considered significantly differentially expressed if the adjusted p-value is less than or equal to 0.05, and the absolute logFC is greater than or equal to a 1.5 fold change. Venn diagrams were made using the matplotlib library venn, and upset plots were made using UpSetPlot in python [71].

<u>Human:</u> We tested gene expression changes in the human dataset using the voom implementation called dream, because of the presence of multiple samples from some individuals [72]. Data was prepared as above for the mouse data. Briefly, expression counts in each replicate, for each cell type with at least 15 cells were pseudobulked, then individual comparisons were subset for downstream analysis. First we removed outlier samples with a log library size 3X median absolute deviation (MAD) below the median, as library size seemed to be one of the strongest drivers of variability. Next, for each comparison, the counts matrices were trimmed using the edgeR filterByExpr function, using a design matrix that accounted for donor age, sex, and post mortem interval, and keep.lib.sizes set to FALSE, then normalization factors were calculated using the TMM method. For dream, we tested several design matrices, but did not find that variance was largely contributed by donor age, sex, or post-mortem interval, but is mainly predicted by donor effect. We therefore ran voomWithDreamWeights on the counts matrix by comparing the group (lesion type) and only using the donor as a fixed effect variable. DEGs were calculated between MOL_G in the lesion and MOL_C in the control WM. We did not include MOL_F because very few cells per patient sample were present, though MOL_F and G are likely similar transcriptional states and share many marker genes. Plots were similarly made with matplotlib or UpSetPlot in python.

Gene set enrichment analysis:

Fast gene set enrichment analysis (fgsea) was implemented to score pathway enrichment [73]. All significant DEGs (adjusted p-value less than or equal to 0.05, and absolute logFC greater than or equal to a 1.5 fold change) were input to the analysis, using the logFC as the score. The msigdbr was used to retrieve the mouse or human hallmark gene set as the input pathway [23]. The output was filtered using minSzie and maxSize to include only pathways with more than 15 genes but less than 500 to avoid noise, then filtered for terms with adjusted p-value less than 0.05. Dotplots were made by combining the output tables of the significant terms, and plotted using ggplot2.

Gene set scoring:

The following gene lists were used to calculate scores using scanpy functions [8,37,74]. Mouse nomenclature was interchanged by keeping only the first capital letter:

Cell cycle scoring: We downloaded the Regev lab cell cycle genes, as used in Seurat and the scverse tutorials. For the mouse calculations, we converted the genes to mouse nomenclature by changing the names from human to mouse formatting (ie. GENE to Gene).

Pandey: GADD45B, CAMK2D, CNTN1, VIM, NUPR1, SLC3A2, FOS, JUNB, JUN, DDIT3, KLK6, CD63, HMGB2, DYNLL1, CSF1, B2M, FBL, KLF6

IRAS1: CCL2, CCL20, CSF3, CXCL1, CXCL2, CXCL3, CXCL5, FTH1, PDCD5, PDLIM4, PITX3, PSME2, SERPINE1, SOD2, TNFAIP6

IRAS2: B2M, CALR, CARHSP1, CCL2, CD47, FTH1, ICAM1, LAP3, NFKBIA, NNMT, PDCD5, PFN1, PKM, PRELID1, PSME2, SOD2, TMSB10, VCAM1

Common innate immunity: IFI47, IIGP1, MDFIC, TOP2A, DYNLT1A, LGALS1, CXCL10, OASL1, ZBP1, DHX58, TRIM25, RSAD2, SLFN9, OAS1A, IFITM2, BCL3, IFI27L2A, OASL2, IRF7, IFIT1, BST2, ISG15, IFIT3B, IFIT3, IFITM3, RTP4

LPC signature: 'KLK8', 'BCAN', 'STEAP3', 'SOX5', 'NRCAM', 'PLVAP', 'CHL1', 'NAV3', 'FGFRL1', 'SEMA4F', 'TSPAN4', 'AMOTL1', 'IFIT1', 'OASL2', 'PIEZO2', 'ANXA2', 'SOCS3', 'GBP3', 'MMD2', 'USP18', 'NPB', 'FAM124A', 'DDX60', 'BCL3', 'RALGPS2', 'PDZRN4', 'APBA2', 'IFI27L2A', 'C4B', 'NNAT', 'SNX10', 'RASIP1', 'CACHD1', 'MASP1', 'PARP3', 'KCNIP1', 'PROS1'

Cuprizone signature: 'SYT4', 'NUPR1','TENM4', 'TRIB3', 'MOXD1', 'GDF15', 'SLC7A5', 'CYB5R2', 'SOX2', 'DNAL4', 'SMPD3', 'DDIT3', 'ATF4', 'EDA2R', 'COX20', 'CEBPG', 'CDH6', 'CHCHD10', 'FRMD4A', 'CD44', 'VEGFA', 'XYLT1', 'PTCH1', 'SMN1'

Meta-Neighborhood analysis:

To determine cluster similarity in our cross-species analysis, we used MetaNeighbor [75]. The merged human-mouse was subset to the HVGs used for integration, and passed to MetaNeighborUS using the fast version, with organism as study_id, and the broad cell type annotation as the ct_col variable.

References:

1. Duncan, I. D. *et al.* The adult oligodendrocyte can participate in remyelination. *Proc. Natl. Acad. Sci.* **115**, E11807–E11816 (2018).
2. Della-Flora Nunes, G. *et al.* Incomplete remyelination via therapeutically enhanced oligodendrogenesis is sufficient to recover visual cortical function. *Nat. Commun.* **16**, 732 (2025).
3. Plemel, J. R. *et al.* Mechanisms of lysophosphatidylcholine-induced demyelination: A primary lipid disrupting myelinopathy. *Glia* **66**, 327–347 (2018).
4. Zirngibl, M., Assinck, P., Sizov, A., Caprariello, A. V. & Plemel, J. R. Oligodendrocyte death and myelin loss in the cuprizone model: an updated overview of the intrinsic and extrinsic causes of cuprizone demyelination. *Mol. Neurodegener.* **17**, 34 (2022).
5. Palavra, F. *et al.* Defining milestones for the study of remyelination using the cuprizone mouse model: How early is early? *Mult. Scler. Relat. Disord.* **63**, 103886 (2022).
6. Leo, H. & Kipp, M. Remyelination in Multiple Sclerosis: Findings in the Cuprizone Model. *Int. J. Mol. Sci.* **23**, 16093 (2022).
7. Shen, K. *et al.* Multiple sclerosis risk gene Mertk is required for microglial activation and subsequent remyelination. *Cell Rep.* **34**, 108835 (2021).
8. Pandey, S. *et al.* Disease-associated oligodendrocyte responses across neurodegenerative diseases. *Cell Rep.* **40**, 111189 (2022).
9. Hou, J. *et al.* Transcriptomic atlas and interaction networks of brain cells in mouse CNS demyelination and remyelination. *Cell Rep.* **42**, 112293 (2023).
10. Jäkel, S. *et al.* Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).
11. Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75–82 (2019).
12. Absinta, M. *et al.* A lymphocyte-microglia-astrocyte axis in chronic active multiple sclerosis. *Nature* **597**, 709–714 (2021).
13. Macnair, W. *et al.* snRNA-seq stratifies multiple sclerosis patients into distinct white matter glial responses. (2025).
14. Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75–82 (2019).
15. Lerma-Martin, C. *et al.* Cell type mapping reveals tissue niches and interactions in subcortical multiple sclerosis lesions. *Nat. Neurosci.* **27**, 2354–2365 (2024).

16. Kukanja, P. *et al.* Cellular architecture of evolving neuroinflammatory lesions and multiple sclerosis pathology. *Cell* **0**, (2024).
17. Castillo-Rodriguez, M. de los A. *et al.* Astroglial and oligodendroglial markers in the cuprizone animal model for de- and remyelination. *Histochem. Cell Biol.* **158**, 15–38 (2022).
18. Molina-Gonzalez, I. *et al.* Astrocyte-oligodendrocyte interaction regulates central nervous system regeneration. *Nat. Commun.* **14**, 3372 (2023).
19. Wolf, F. A. *et al.* PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
20. Valihrach, L. *et al.* Recent advances in deciphering oligodendrocyte heterogeneity with single-cell transcriptomics. *Front. Cell. Neurosci.* **16**, (2022).
21. Floriddia, E. M. *et al.* Distinct oligodendrocyte populations have spatial preference and different responses to spinal cord injury. *Nat. Commun.* **11**, 5860 (2020).
22. Park, H. *et al.* Single-cell RNA-sequencing identifies disease-associated oligodendrocytes in male APP NL-G-F and 5XFAD mice. *Nat. Commun.* **14**, 802 (2023).
23. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. (2015).
24. Praet, J., Guglielmetti, C., Berneman, Z., Van der Linden, A. & Ponsaerts, P. Cellular and molecular neuropathology of the cuprizone mouse model: Clinical relevance for multiple sclerosis. *Neurosci. Biobehav. Rev.* **47**, 485–505 (2014).
25. Stone, S. & Lin, W. The unfolded protein response in multiple sclerosis. *Front. Neurosci.* **9**, 264 (2015).
26. Emery, B. *et al.* Suppressor of cytokine signaling 3 limits protection of leukemia inhibitory factor receptor signaling against central demyelination. *Proc. Natl. Acad. Sci.* **103**, 7859–7864 (2006).
27. Kirby, L. *et al.* Oligodendrocyte precursor cells present antigen and are cytotoxic targets in inflammatory demyelination. *Nat. Commun.* **10**, 3887 (2019).
28. Falcão, A. M. *et al.* Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. *Nat. Med.* **24**, 1837–1844 (2019).
29. Hamanaka, G. *et al.* Myelination- and migration-associated genes are downregulated after phagocytosis in cultured oligodendrocyte precursor cells. *J. Neurochem.* **167**, 571–581 (2023).
30. Way, S. W. *et al.* Pharmaceutical integrated stress response enhancement protects oligodendrocytes and provides a potential multiple sclerosis therapeutic. *Nat. Commun.* **6**, 6532 (2015).

31. Lou, J. *et al.* SNX10 promotes phagosome maturation in macrophages and protects mice against Listeria monocytogenes infection. *Oncotarget* **8**, 53935–53947 (2017).
32. Gimple, R. C. *et al.* Sorting nexin 10 sustains PDGF receptor signaling in glioblastoma stem cells via endosomal protein sorting. *JCI Insight* **8**, e158077 (2023).
33. Tuohy, T. M. F. *et al.* CD44 overexpression by oligodendrocytes: A novel mouse model of inflammation-independent demyelination and dysmyelination. *Glia* **47**, 335–345 (2004).
34. Pruvost, M. *et al.* The stability of the myelinating oligodendrocyte transcriptome is regulated by the nuclear lamina. *Cell Rep.* **42**, 112848 (2023).
35. Filipek, A. & Leśniak, W. Current view on cellular function of S100A6 and its ligands, CacyBP/SIP and Sgt1. *Postepy Biochem.* **64**, 242–252 (2018).
36. Paolicelli, R. C. *et al.* Microglia states and nomenclature: A field at its crossroads. *Neuron* **110**, 3458–3483 (2022).
37. Koupourtidou, C. *et al.* Shared inflammatory glial cell signature after stab wound injury, revealed by spatial, temporal, and cell-type-specific profiling of the murine cerebral cortex. *Nat. Commun.* **15**, 2866 (2024).
38. March-Diaz, R. *et al.* Hypoxia compromises the mitochondrial metabolism of Alzheimer's disease microglia via HIF1. *Nat. Aging* **1**, 385–399 (2021).
39. Zhou, Z. *et al.* SASH1 contributes to glial cell migration in the early development of the central nervous system. *Dev. Biol.* **504**, 49–57 (2023).
40. Liu, S. *et al.* Depletion of SASH1, an astrocyte differentiation-related gene, contributes to functional recovery in spinal cord injury. *CNS Neurosci. Ther.* **29**, 228–238 (2022).
41. Jaufmann, J. *et al.* The emerging and diverse roles of the SLy/SASH1-protein family in health and disease—Overview of three multifunctional proteins. *FASEB J.* **35**, e21470 (2021).
42. Huntula, S., Saegusa, H., Wang, X., Zong, S. & Tanabe, T. Involvement of N-type Ca2+ channel in microglial activation and its implications to aging-induced exaggerated cytokine response. *Cell Calcium* **82**, 102059 (2019).
43. van der Poel, M. *et al.* Transcriptional profiling of human microglia reveals grey–white matter heterogeneity and multiple sclerosis-associated changes. *Nat. Commun.* **10**, 1139 (2019).
44. Liu, J. *et al.* NUPR1 is a critical repressor of ferroptosis. *Nat. Commun.* **12**, 647 (2021).
45. Girolamo, F., Coppola, C., Ribatti, D. & Trojano, M. Angiogenesis in multiple sclerosis and experimental autoimmune encephalomyelitis. *Acta Neuropathol. Commun.* **2**, 84 (2014).
46. Lin, W. Neuroprotective effects of vascular endothelial growth factor A in the experimental autoimmune encephalomyelitis model of multiple sclerosis. *Neural Regen. Res.* **12**, 70–71

(2017).
47. Geirsdottir, L. *et al.* Cross-Species Single-Cell Analysis Reveals Divergence of the Primate Microglia Program. *Cell* **179**, 1609-1622.e16 (2019).
48. Bogie, J. F. J. *et al.* Scavenger receptor collectin placenta 1 is a novel receptor involved in the uptake of myelin by phagocytes. *Sci. Rep.* **7**, 44794 (2017).
49. Thines, L., Roushar, F. J., Hedman, A. C. & Sacks, D. B. The IQGAP scaffolds: Critical nodes bridging receptor activation to cellular signaling. *J. Cell Biol.* **222**, e202205062 (2023).
50. Hrovatin, K. *et al.* Integrating single-cell RNA-seq datasets with substantial batch effects. *bioRxiv* 2023.11.03.565463 (2024) doi:10.1101/2023.11.03.565463.
51. Hooijmans, C. R. *et al.* Remyelination promoting therapies in multiple sclerosis animal models: a systematic review and meta-analysis. *Sci. Rep.* **9**, 822 (2019).
52. Madsen, P. M. *et al.* Oligodendroglial TNFR2 Mediates Membrane TNF-Dependent Repair in Experimental Autoimmune Encephalomyelitis by Promoting Oligodendrocyte Differentiation and Remyelination. *J. Neurosci.* **36**, 5128–5143 (2016).
53. Zhan, J. *et al.* The Cuprizone Model: Dos and Do Nots. *Cells* **9**, 843 (2020).
54. Haroon, A., Seerapu, H., Fang, L.-P., Weß, J. H. & Bai, X. Unlocking the Potential: immune functions of oligodendrocyte precursor cells. *Front. Immunol.* **15**, 1425706 (2024).
55. Miron, V. E. *et al.* M2 microglia and macrophages drive oligodendrocyte differentiation during CNS remyelination. *Nat. Neurosci.* **16**, 1211–1218 (2013).
56. Lampron, A. *et al.* Inefficient clearance of myelin debris by microglia impairs remyelinating processes. *J. Exp. Med.* **212**, 481–495 (2015).
57. Lloyd, A. F. & Miron, V. E. The pro-remyelination properties of microglia in the central nervous system. *Nat. Rev. Neurol.* **15**, 447–458 (2019).
58. Réu, P. *et al.* The Lifespan and Turnover of Microglia in the Human Brain. *Cell Rep.* **20**, 779–784 (2017).
59. Mathys, H. *et al.* Temporal Tracking of Microglia Activation in Neurodegeneration at Single-Cell Resolution. *Cell Rep.* **21**, 366–380 (2017).
60. Sadick, J. S. *et al.* Astrocytes and oligodendrocytes undergo subtype-specific transcriptional changes in Alzheimer's disease. *Neuron* **110**, 1788-1805.e10 (2022).
61. Skripuletz, T. *et al.* Astrocytes regulate myelin clearance through recruitment of microglia during cuprizone-induced demyelination. *Brain* **136**, 147–167 (2013).
62. Fleming, S. J. *et al.* Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* **20**, 1323–1335 (2023).

63. Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nat. Rev. Genet.* **24**, 550–572 (2023).
64. Bernstein, N. J. *et al.* Solo: Doublet Identification in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell Syst.* **11**, 95-101.e5 (2020).
65. Macnair, W. *et al.* Single nuclei RNAseq stratifies multiple sclerosis patients into distinct white matter glial responses. 2022.04.06.487263 Preprint at https://doi.org/10.1101/2022.04.06.487263 (2023).
66. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
67. Gayoso, A. *et al.* A Python library for probabilistic analysis of single-cell omics data. *Nat. Biotechnol.* **40**, 163–166 (2022).
68. Hrovatin, K. *et al.* Considerations for building and using integrated single-cell atlases. *Nat. Methods* 1–17 (2024) doi:10.1038/s41592-024-02532-y.
69. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).
70. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
71. Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R. & Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992 (2014).
72. Hoffman, G. E. & Roussos, P. Dream: powerful differential expression analysis for repeated measures designs. *Bioinformatics* **37**, 192–201 (2020).
73. Korotkevich, G. *et al.* Fast gene set enrichment analysis. 060012 Preprint at https://doi.org/10.1101/060012 (2021).
74. Leng, K. *et al.* CRISPRi screens in human iPSC-derived astrocytes elucidate regulators of distinct inflammatory reactive states. *Nat. Neurosci.* **25**, 1528–1542 (2022).
75. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat. Commun.* **9**, 884 (2018).