# Disease-associated oligodendrocyte responses across neurodegenerative diseases

Graphical abstract

Authors

Shristi Pandey, Kimberle Shen,
Seung-Hye Lee, ..., Christopher J. Bohlen,
Tracy J. Yuen, Brad A. Friedman

Correspondence

pandey.shristi@gene.com (S.P.),
tjyuen@gmail.com (T.J.Y.),
friedman.brad@gene.com (B.A.F.)

In brief

Using an integrative analysis of single-cell
RNA-seq datasets across disease
models of AD and MS, Pandey et al.
identified three distinct activation states
of oligodendrocytes. Human MS
oligodendrocytes share activation genes
with mouse models, whereas the human
AD response is largely distinct.

Highlights

- Oligodendrocyte lineage meta-analysis of AD and MS mouse model scRNA-seq datasets

- We find three distinct disease-associated oligodendrocytes (DAOs) across disease models

- DAOs are established outside the demyelinated lesion and repopulate it during remyelination

- The human MS oligo response shows overlap but the human AD response is distinct from DAOs

CellPress

Cell Reports

Resource

Disease-associated oligodendrocyte responses across neurodegenerative diseases

Shristi Pandey,[1,6,*] Kimberle Shen,[2,5] Seung-Hye Lee,[2] Yun-An A. Shen,[2] Yuanyuan Wang,[2] Marcos Otero-García,[3] Natalya Kotova,[3] Stephen T. Vito,[2] Benjamin I. Laufer,[1] Dwight F. Newton,[1,4] Mitchell G. Rezzonico,[1] Jesse E. Hanson,[2] Joshua S. Kaminker,[1] Christopher J. Bohlen,[2] Tracy J. Yuen,[2,*] and Brad A. Friedman[1,*]

[1]Department of OMNI Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
[2]Department of Neuroscience, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
[3]Rebus Biosystems, Inc., Santa Clara, CA, USA
[4]Roche Global IT Solution Centre, Hoffman-La Roche Canada, 7070 Mississauga Road, Mississauga, ON, Canada
[5]Present address: Genome Institute of Singapore, Agency for Science, Technology, and Research, Singapore, Singapore
[6]Lead contact
*Correspondence: pandey.shristi@gene.com (S.P.), tjyuen@gmail.com (T.J.Y.), friedman.brad@gene.com (B.A.F.)
https://doi.org/10.1016/j.celrep.2022.111189

SUMMARY

Oligodendrocyte dysfunction has been implicated in the pathogenesis of neurodegenerative diseases, so understanding oligodendrocyte activation states would shed light on disease processes. We identify three distinct activation states of oligodendrocytes from single-cell RNA sequencing (RNA-seq) of mouse models of Alzheimer's disease (AD) and multiple sclerosis (MS): DA1 (disease-associated1, associated with immunogenic genes), DA2 (disease-associated2, associated with genes influencing survival), and IFN (associated with interferon response genes). Spatial analysis of disease-associated oligodendrocytes (DAOs) in the cuprizone model reveals that DA1 and DA2 are established outside of the lesion area during demyelination and that DA1 repopulates the lesion during remyelination. Independent meta-analysis of human single-nucleus RNA-seq datasets reveals that the transcriptional responses of MS oligodendrocytes share features with mouse models. In contrast, the oligodendrocyte activation signature observed in human AD is largely distinct from those observed in mice. This catalog of oligodendrocyte activation states (http://research-pub.gene.com/OligoLandscape/) will be important to understand disease progression and develop therapeutic interventions.

INTRODUCTION
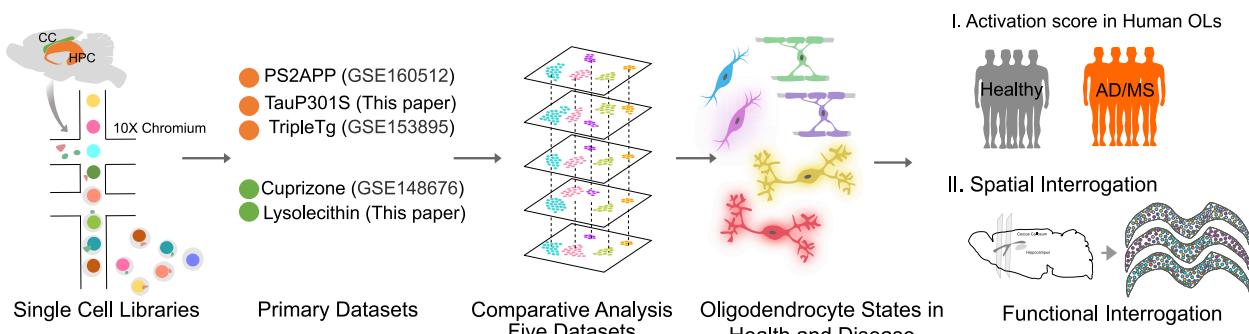
Oligodendrocytes (OLs) generate myelin, enabling fast saltatory propagation of axon potentials (Simons and Nave, 2016), and provide metabolic support to axons (Fünfschilling et al., 2012; Griffiths et al., 1998, 2021c). Selective loss of myelin and OLs contributes to the pathogenesis of progressive neurodegenerative diseases such as multiple sclerosis (MS) (Trapp and Nave, 2008) and leukodystrophies (Knaap and Bugiani, 2017). Similarly, white matter degeneration and myelin loss have been observed in other neurodegenerative conditions such as Alzheimer's disease (AD) (Caso et al., 2015). Age-related myelin breakdown has been postulated to presage AD (Bartzokis, 2011), and the pattern of neurofibrillary changes also bears an inverse resemblance to the pattern of myelination (Braak and Braak, 1996).

While other glial cells, such as microglia, have been extensively characterized in the context of neurodegeneration, OLs have received far less attention. Recent studies have uncovered the molecular heterogeneity of OLs in various brain regions during development (Marques et al., 2016) and in diseases such as MS (Absinta et al., 2021; Jäkel et al., 2019; Schirmer et al., 2019) and AD (Cain et al., 2020; Grubman et al., 2019; Lau et al., 2020; Leng et al., 2021; Mathys et al., 2019; Zhou et al., 2020). Transcriptional responses of OLs to pathological conditions have also been observed in animal models of neurodegenerative disease (Falcão et al., 2018; Floridia et al., 2020; Lee et al., 2021a; Shen et al., 2021; Zhou et al., 2020). These heterogeneous populations of OLs may have differing roles in responding to demyelination and other disease pathologies. Indeed, developmentally distinct populations of OL precursor cells (OPCs) were recently shown to generate OL-lineage cells with differing abilities to respond to demyelination (Crawford et al., 2016). Furthermore, OPCs have been shown to have roles beyond those related to myelination, including regulation of angiogenesis in the normal postnatal brain (Yuen et al., 2014) and antigen presentation and phagocytosis in mouse models of MS (Falcão et al., 2018; Kirby et al., 2019). While these studies have sought to define molecular signatures of OL subtypes and characterize their transcriptional responses in diseased states and models, a systematic understanding of how these transcriptional profiles interrelate with one another across these disease states and models is lacking. More importantly, it is unclear which aspects of OL responses in human disease are recapitulated in animal models.

A

I. Activation score in Human OLs

Healthy   AD/MS

II. Spatial Interrogation

Single Cell Libraries   Primary Datasets   Comparative Analysis Five Datasets   Oligodendrocyte States in Health and Disease   Functional Interrogation

PS2APP (GSE160512)
TauP301S (This paper)
TripleTg (GSE153895)
Cuprizone (GSE148676)
Lysolecithin (This paper)

B

Datasets
PS2APP
TauP301S
TripleTg
Cuprizone
Lysolecithin

n=5 datasets
n=55 mice
n=118,279 cells

C

MOL2_IFN
MOL2_DA1
Mature Oligodendrocyte2 (MOL2)
Newly formed Oligodendrocytes2(NFOL2)
Myelin Forming Oligodendrocytes (MFOL)
Mature Oligodendrocyte1(MOL1)
MOL5/6_DA2
Mature Oligodendrocytes5/6 (MOL5/6)
treatment MOL5/6 (tMOL5/6)
Interferon-related MOL5/6 (MOL5/6_IFN)
MOL5/6_DA1
Newly formed Oligodendrocytes1(NFOL1)
DA2 like-OPC
Committed OL Precursors(COP)
Interferon-related OPC (OPC_IFN)
Oligodendrocyte Precursor Cells(OPC)
Cycling OPC (cycOPC)

0 0.25 0.5 0.75 1.0
Percent of Cluster

Datasets: PS2APP   TauP301S   TripleTg   Cuprizone   Lysolecithin

D

(legend on next page)

Here, we performed an integrative single-cell analysis of OL-lineage cells across a wide range of AD and MS models. We found three disease-associated OLs (DAOs) in mouse models: a DA1 state, associated with elevated expression of inflammatory genes; a DA2 state, associated with elevated expression of genes that influence OL survival; and an interferon-related IFN state. We validated these states using *in situ* hybridization and also observed their distribution relative to cuprizone-induced lesions. Finally, we analyzed human single-nucleus RNA-sequencing (snRNA-seq) datasets from three MS studies and six AD studies to understand which aspects of the DAO gene program are conserved in human disease. While MS patient OLs shared features of the DAO gene program, AD OLs were largely distinct. This understanding of DAOs and their respective gene programs will be a valuable resource to gain mechanistic insights and to target OL-lineage cells for regeneration and repair of the nervous system in various neurodegenerative diseases.

RESULTS

Multi-dataset integration reveals OL transcriptional states in AD models and remyelination models of MS

To understand the transcriptional responses of OLs in diverse AD and MS models, we integrated OL-lineage cells from five primary datasets spanning three AD models and two remyelination models of MS (Figures 1A and 1B). Among the AD models, we included OL-lineage cells from well-characterized amyloidosis models such as the PS2APP mice (Ozmen et al., 2008; Richards et al., 2003), tauopathy models such as mice harboring the TauP301S (Yoshiyama et al., 2007) or the TauP301L (Götz et al., 2001) mutation, and TauPS2APP animals bearing combined amyloid and tau pathology (Grueninger et al., 2010; Lee et al., 2021a). Among the remyelination models, we included mice that underwent cuprizone- or lysolecithin-induced demyelination. Cuprizone exposure for 4 weeks leads to OL cell death and robust demyelination followed by spontaneous remyelination after cuprizone withdrawal (Matsushima and Morell, 2001). Lysolecithin injection is a focal and acute injury resulting in demyelination at the injection site, followed by predictable kinetics of spontaneous remyelination, mostly complete by 28 days (Gensert and Goldman, 1997; Miron et al., 2013). These different models were collected as separate datasets, except the TauP301L and TauPS2APP models, which were part of a single dataset, collectively referred to here as the TripleTg dataset (Figure 1A). The TripleTg dataset also contained Trem2^{KO} animals harboring tau and amyloid pathology (TauPSAPP, Trem2^{KO}), allowing us to understand if aspects of the DAO response could be dependent on *Trem2* activation. The datasets representing AD models were generated from the hippocampus, whereas those for remyelination models were generated from the corpus callosum. These datasets formed part of studies that were published in our previous work, except the lysolecithin and the TauP301S model, which represent unpublished data (Table S1) (Lee et al., 2021a; Shen et al., 2021). To identify shared and distinct OL transcriptional states from these diverse datasets, we employed Seurat's data integration pipeline (Stuart et al., 2019). We first identified the OL-lineage cells from each dataset based on combinatorial expression of OL-lineage markers such as *Pdgfra*, *Vcan*, *Sox10*, *Mog*, *Myrf*, *Mag*, and *Plp1*. Each dataset was filtered to remove low-quality cells with less than 400 unique molecular identifiers (UMIs) or greater than 5% mitochondrially encoded genes (Figure S1A). These OLs were then integrated either along datasets or in sample batches based on whether batch effects were observed within each dataset. After integration, the cells were clustered and labeled based on marker genes conserved across all datasets (Figure S1B). No batch effects were observed in each dataset post-integration (Figure S1C, top).
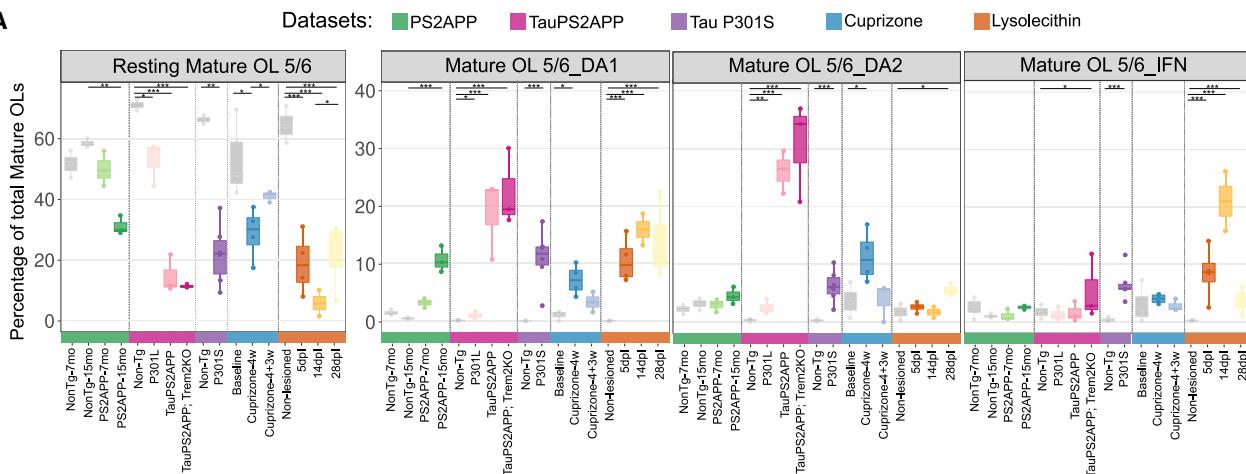
We identified a total of 17 clusters in our integrated analysis (Figure 1C) of 118,279 cells from the five primary datasets (Figure 1B). Six clusters represented OPCs and committed OL precursors (COPs), identified based on canonical markers (Marques et al., 2016, 2018). Three clusters represented immature OL subtypes such as newly formed OLs (NFOL1 and NFOL2) and myelin-forming OLs (MFOLs), whereas nine others represented different mature OL (MOL) subtypes and their disease-associated transcriptional states (Figures 1C and 1D). Of the six mature resting subtypes identified in previous studies, we were able to recover three MOL subtypes, namely MOL1, MOL2, and MOL5/6, in our analysis. These MOL subtypes expressed canonical subtype-specific markers, such as *Egr1*, *Btg2*, *Klf4*, and *Arc* in MOL1; *Hopx*, *Plin3*, *Klk6*, and *S100b* in MOL2; and *Ptgds*, *Il33*, and *Opalin* in MOL5/6 (Figures 1D and S1D). Of these, MOL1 may represent a dissociation-induced state similar to what has previously been described for microglia (Haimon et al., 2018), as it is characterized by immediate-early gene expression and detected in datasets that were prepared for single-cell RNA-seq (scRNA-seq) without the addition of actinomycin D before cell capture (Figure S1E, PS2APP, cuprizone, and lysolecithin) (Wu et al., 2017).

Each mature subtype, MOL2 and MOL5/6, had disease-associated subclusters that were so named because they were enriched in the transgenic or treated animals in each dataset (Figure S1C, bottom). While these DAOs expressed a
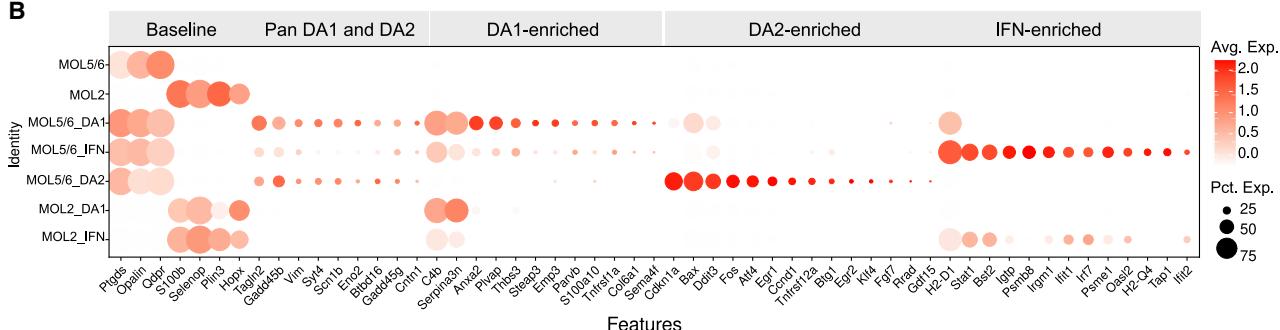

Figure 1. Multi-dataset integration of single-cell RNA-seq profiles of OL-lineage cells across brain regions and mouse models

(A) Workflow for single-cell dataset integration of MS and AD mouse models.

(B) Uniform manifold approximation and projection (UMAP) plot of five independent datasets from mouse models of neurodegeneration integrated into a single analysis using the Seurat dataset integration pipeline.

(C) UMAP plot highlighting the major resting and disease-associated oligodendrocytes (DAOs). Cluster labels are represented in the form of a dendrogram that captures global intercluster transcriptional relationships. Dendrogram was built by performing hierarchical clustering on the average gene-expression profiles for each cluster, restricting to highly variable genes in the data. The percentage of cells from each dataset contributing to each cluster is highlighted in the bar plot on the right.

(D) Dot plot of representative marker genes enriched in baseline OL-lineage subtypes and DAOs. MOL subtypes and transcriptional states are plotted on the y axis. Genes are plotted on the x axis. The color scale represents the average expression of a given gene in the cluster, and the size of the dot represents the percentage of cells in the cluster that express a given gene. See also Figure S1 and Table S1.

A Datasets: ■ PS2APP ■ TauPS2APP ■ Tau P301S ■ Cuprizone ■ Lysolecithin

B

C

D

E

● PS2APP   ● TauPS2APP   ● Tau P301S   ● Cuprizone   ● Lysolecithin

F MOL5/6_DA1

G MOL5/6_DA2

H MOL5/6_IFN
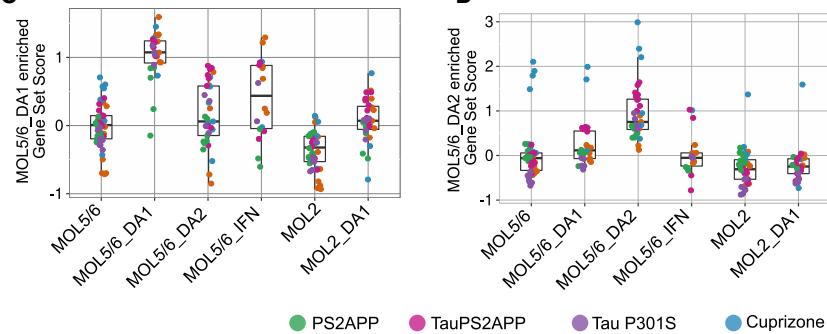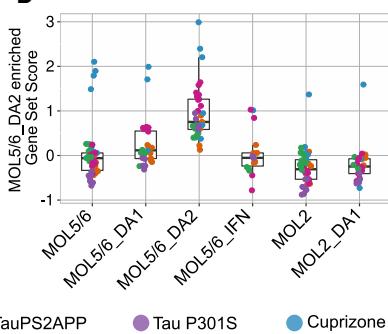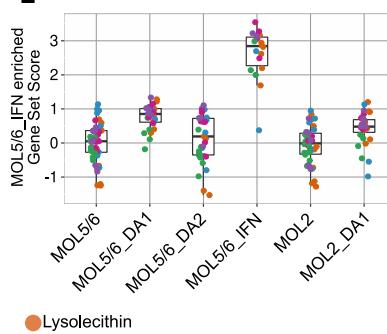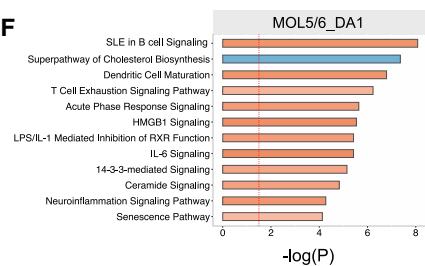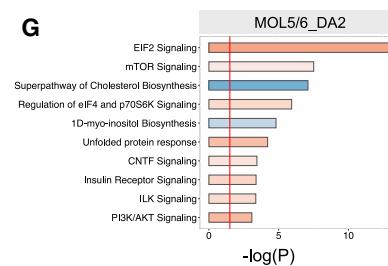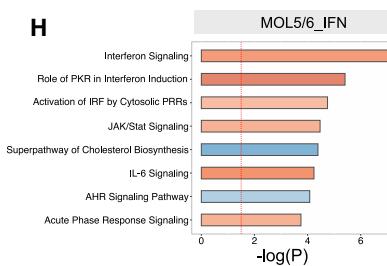
Low Z-score High

(legend on next page)

majority of the baseline MOL-subtype markers identified by previous work, they also expressed additional genes that were either very low or absent in healthy controls (Figure 1D). In the next section, we describe these DAOs and their respective gene programs.

Three shared disease-associated states identified in MOL5/6 oligodendrocytes across multiple AD and MS models

Of the three MOLs observed in our data, MOL5/6 is the most abundant subtype, comprising 50%–70% of MOLs in datasets derived from both the hippocampus and the corpus callosum (Figure 2A, resting MOL5/6). We observed three disease-associated states of MOL5/6, which we termed MOL5/6_DA1 (with high expression of *C4b*, *Serpina3n*), MOL5/6_DA2 (with high expression of *Cdkn1a*, *Tnfrsf12a*), and MOL5/6_IFN (with high expression of interferon-stimulated genes such as members of the *Ifit* and *Oas* families) (Figure 1D). Concurrent with the significant decrease in the proportion of resting MOL5/6 in treatment groups across the five datasets, there is a varying increase in the fraction of cells in disease-associated MOL5/6 states in all models included in our analysis (Figure 2A).

Among the AD models, we observed a significant increase in the proportion of MOL5/6_DA1 state across all models except the young 7-month PS2APP and the P301L animals. This increase in proportion was highest in the TauPS2APP animals, followed by the TauP301S and old 15-month PS2APP animals. Unlike the MOL5/6_DA1 state, the MOL5/6_DA2 state did not increase in old PS2APP animals but increased substantially in the TauPS2APP and less so in the TauP301S animals (Figure 2A). Neither of the two states was dependent on *Trem2* as seen by their near comparable activation in TauPS2APP and TauPS2APP; Trem2[KO] animals. This is in line with the *Trem2* independence reported in the induction of *C4b* and *Serpina3n* in cortical OLs in the 15-month old 5XFAD animals (Zhou et al., 2020).

Among the remyelination models, both cuprizone- and lysolecithin-treated animals undergo a predictable decrease in the number of OLs at the 4-week time point and 5 days post-lesion (dpl), respectively. The small number of remaining OLs at these peak demyelination time points shifts to the aforementioned subtypes. Among these, MOL5/6_DA1 increased in proportion in both models and did not return to baseline in any of the later time points examined (Figure 2A). Interestingly, we found a striking difference in the timing of appearance of the MOL5/6_DA2 state in the two remyelination models. In cuprizone animals, the DA2 state markedly increased at peak demyelination (4 weeks), yet it appeared in the lysolecithin animals only at 28 dpl, when remyelination is nearly complete (Figure 2A). This

might reflect the difference in remyelination kinetics in the two models: while the cuprizone-treated animals do not undergo substantial remyelination until cuprizone is withdrawn at 4 weeks, the lysolecithin-treated animals are known to undergo active remyelination at 10–14 dpl (Gensert and Goldman, 1997; Miron et al., 2013). Alternatively, it is possible that DA2 cells do appear at earlier time points in the lysolecithin model, but outside of the lesion area dissected for this analysis. The lysolecithin-treated animals also display a marked upregulation of the MOL5/6_IFN in the earlier time points at 5 and 14 dpl. Strikingly, the transcriptional state shift to DA1 is persistent at recovery time points (4 + 3 weeks and 28 dpl) in both models.

Characterization of gene sets and pathways associated with disease-associated MOL5/6 subtypes

To identify broad gene signatures associated with each DAO cluster, we generated pseudo-bulk profiles of resting MOL5/6 and an average of all disease-associated MOL5/6 clusters and performed differential expression analysis between the two compartments (Figures S2A and S3A). To identify the gene signatures associated with each cluster of DAOs (example genes in Figure 2B), we aggregated single cells belonging to MOL5/6, MOL5/6_DA1, MOL5/6_DA2, and MOL5/6_IFN into separate pseudo-bulk profiles (Figure S2B). Comparing each disease-associated MOL5/6 to resting MOL5/6, we obtained 128, 107, and 78 differentially expressed (DE) genes unique to MOL5/6_DA1 (Figure 2C), MOL5/6_DA2 (Figure 2D), and MOL5/6_IFN (Figure 2E), and 303 genes shared among different combination of states (Figure S2C). Given that the majority of DE genes were expressed in a gradient across the transcriptional states (Figures 2B and S2D) at varying fold changes, we classified genes as being enriched in a particular state if their fold change was the maximum in that state. Genes were then classified as enriched in multiple states if the fold change of that gene was comparable in another DAO cluster (STAR Methods). Example genes from pan DA1 and DA2 (Figure S2E), DA1-specific (Figure S2F), DA2-specific (Figure S2G), and IFN-specific (Figure S2H) gene sets highlight their expression across all the datasets incorporated in this study.

Pathway analysis of the DE genes in MOL5/6_DA1 showed an enrichment in immune-related pathways, including genes involved in inflammatory response, such as *Tnfrsf1a*, *C4b*, *Il1b*, *Hmox1*, *Tnf*, and *Serpina3n*, and genes specifically implicated in neuroinflammation, such as *Bace2* and *B2m* (Figures 2B and 2F and S4A). In addition, a number of genes involved in antigen presentation, such as the major histocompatibility complex (MHC) class I and MHC class II molecules, are also upregulated in these OLs (Figure S4A). Together, these results suggest an elevated inflammatory signature in MOL5/6_DA1, indicating

---

Figure 2. MOL5/6 exhibits three distinct activation states

(A) Relative proportions of MOL5/6_DA1, MOL5/6_DA2, and MOL5/6_IFN in each genotype or time point in five datasets represented in the form of a boxplot. Differential abundance statistics: *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001 compared with baseline or non-transgenic condition. FDR, false discovery rate.

(B) Dot plot of marker genes enriched in baseline MOLs, MOL5/6_DA1, MOL5/6_DA2, and MOL5/6_IFN.

(C–E) DE gene set scores of the genes derived by comparing each disease-associated MOL5/6 to baseline MOL5/6: MOL5/6_DA1 (C), MOL5/6_DA2 (D), and MOL5/6_IFN (E). Only unique genes for each state were included.

(F–H) Ingenuity pathway analysis of the DE genes in MOL5/6_DA1 (F), MOL5/6_DA2 (G), and MOL5/6_IFN (H). High Z scores depict predicted activation and low Z scores represent predicted inhibition of respective pathways. All DE genes associated with each state were included. See also Figures S2–S4 and Table S2.

that OLs in disease could be not just the target of a dysregulated immune system but also immunomodulatory themselves.

In contrast, genes upregulated in the MOL5/6_DA2 state belong to pathways that may influence the survival of OLs, such as the EIF2 signaling pathway (*Ccnd1*, *Atf3*, *Bcl2*, *Ddit3*, *Trib3*) and the PI3K/AKT signaling pathway (*Cdkn1a*, *Trp53*, *Itgb2*, *Mdm2*, *Gdf15*) (Ishii et al., 2019; Li et al., 2004) (Figures 2B and 2G). Pathways indicative of cellular stress such as the unfolded protein response are also elevated in these OLs. Similarly, a number of pathways that promote myelination, such as mTOR, insulin receptor, and CNTF signaling, are also upregulated in the DA2 state (Lebrun-Julien et al., 2014; Stankoff et al., 2002; Tyler et al., 2009). Some genes elevated in the DA2 state, such as *Rras1* and *Rras2*, have been directly implicated in driving myelination during development (Sanz-Rodriguez et al., 2018). Together, these results suggest that the DA2 gene programs may influence OL survival as well as myelination.

Last, interferon-related MOL5/6 upregulated genes involved in the interferon response from the *Ifit* and *Oas* families and the *Irf* family of transcription factors, which regulate the expression of interferons (Figure 2H). Antigen presentation genes such as MHC class I and class II genes, proteasomes, and transporters involved in antigen processing, such as *Psmb8*, *Psme1*, *Tap1*, and *Tapbp*, are also robustly upregulated in the MOL5/6_IFN cluster (Figures 2B and S4A). An interferon-related state is also observed among OPCs (Figures S5A–S5C), which upregulate the same signature as MOL5/6_IFN. A similar class of OPCs that upregulated antigen presentation genes was also described in the inflammatory demyelinating adoptive-transfer cuprizone model (Kirby et al., 2019).

While there are marked differences in the pathways upregulated by each disease-associated MOL5/6, we observed a striking similarity in the downregulated genes. All disease-associated clusters of MOL5/6 robustly downregulated cholesterol biosynthetic pathway genes (Figures 2F–2H), such as *Hmgcs1*, *Dhcr7*, and *Acat2*. Each disease-associated state of MOL5/6 is observed in both hippocampus and corpus callosum, indicating that they are not restricted to a single brain region (Figure S4B).

Despite the large differences in pathology across the AD and remyelination models, there is a striking similarity in the transcriptional responses of OLs (Figures S3A and S4B). Similar to switching of homeostatic microglia to disease-associated microglia (DAMs) (Deczkowska et al., 2018; Friedman et al., 2018; Keren-Shaul et al., 2017), our meta-analysis indicates that OLs switch their transcriptional state from homeostatic/resting states to disease-associated states (DAOs), downregulating the expression of resting OL genes such as *Mag* and *Mog* and upregulating the expression of disease-associated genes (Figures 1D and S3A and S4A). Together, these results establish the disease-associated states of MOL5/6 and suggest common activation patterns across diverse models of neurodegeneration.

MOL2 demonstrates an inflammatory DA1 and interferon-related IFN state that is shared with MOL5/6

MOL2 comprised 10%–15% of the total MOLs in both the corpus callosum and the hippocampus and displayed two disease-associated states: *Klk8*, *C4b*, and *Serpina3n*+ MOL2_DA1 and an interferon-related MOL2_IFN (Figure 3A). Similar to the

MOL5/6_DA1 state, MOL2_DA1 was induced in animals with amyloid and tau as well as the TauPS2APP animals (Figure 3B). In lysolecithin-treated animals, MOL2_DA1 increased in proportion during the demyelination phase but decreased during the recovery time points, although not quite returning to baseline (Figure 3B). Cuprizone-treated animals showed similar trends but were not significant in differential abundance analysis.

We also generated a list of DE genes comparing MOL2_DA1 with its resting counterpart MOL2. DE genes between MOL2_DA1 and MOL2 correlated well with the DE genes between MOL5/6_DA1 and MOL5/6, but not well with the DE genes between MOL5/6_DA2 and MOL5/6, suggesting that MOL2_DA1 and MOL5/6_DA1 represent a similar disease-associated state that manifests in both MOL subtypes (Figures 3C and 3D). Only a handful of MOL2_DA1-associated DE genes, such as *Klk8* and *Fxyd7*, are upregulated significantly more in MOL2_DA1 compared with MOL5/6_DA1 (Figure 3C). Pathway analysis of the DE genes enriched in MOL2_DA1 highlights a similar immune signature that was noted in MOL5/6_DA1 (Figure 3E). Similarly, the interferon-related MOL2_IFN represented a small cluster of MOL2 that upregulated the interferon-related gene set similar to MOL5/6_IFN (Figure 3A).

Interestingly, we did not observe a DA2-like state for MOL2, suggesting that the two MOL populations might have differing abilities to respond to disease pathology. However, the lack of a DA2-like state for MOL2 may also be a consequence of lower sampling of this rarer of the two MOL populations. Together, these results establish OLs as dynamic architects during neurodegenerative disease, able to execute three major disease-associated transcriptional states: a DA1, a DA2, and an IFN state. These disease-associated states, when combined with the baseline heterogeneity of MOLs, give OLs a repertoire of transcriptional states that may have varied functional implications for disease progression and outcomes.

DAOs are established outside of microglia-infiltrated lesions and then spread throughout the corpus callosum during remyelination in the cuprizone model

Our scRNA-seq meta-analysis uncovered diverse OL responses in different pathological contexts. Our previous study validated MOL_DA1 and MOL_DA2 states in the amyloid and tau models of AD (Lee et al., 2021a). We sought to further validate these findings and investigate the spatial localization of DAOs with respect to other glial subtypes such as DAMs and disease-associated astrocytes (DAAs) in a remyelination model of MS. To this end, we assayed the expression of candidate DAO markers along with markers of other glial subpopulations in the cuprizone model using multiplexed single-molecule fluorescence *in situ* hybridization (smFISH). We quantified the expression level of 31 RNA transcripts that could collectively identify (1) OPC and OL-lineage states, including resting MOL subtypes and DAO states; (2) homeostatic microglia and DAMs; (3) homeostatic astrocytes and DAAs; (4) endothelial cells; and (5) neurons (Figure 4A). Consistent with the time points assayed by our scRNA-seq data, we collected smFISH at baseline (n = 4), 4 weeks post cuprizone treatment (4 weeks, n = 4), and 4 weeks of cuprizone treatment followed by 3 weeks recovery (4 + 3 weeks, n = 4). Data were collected from the anterior part of
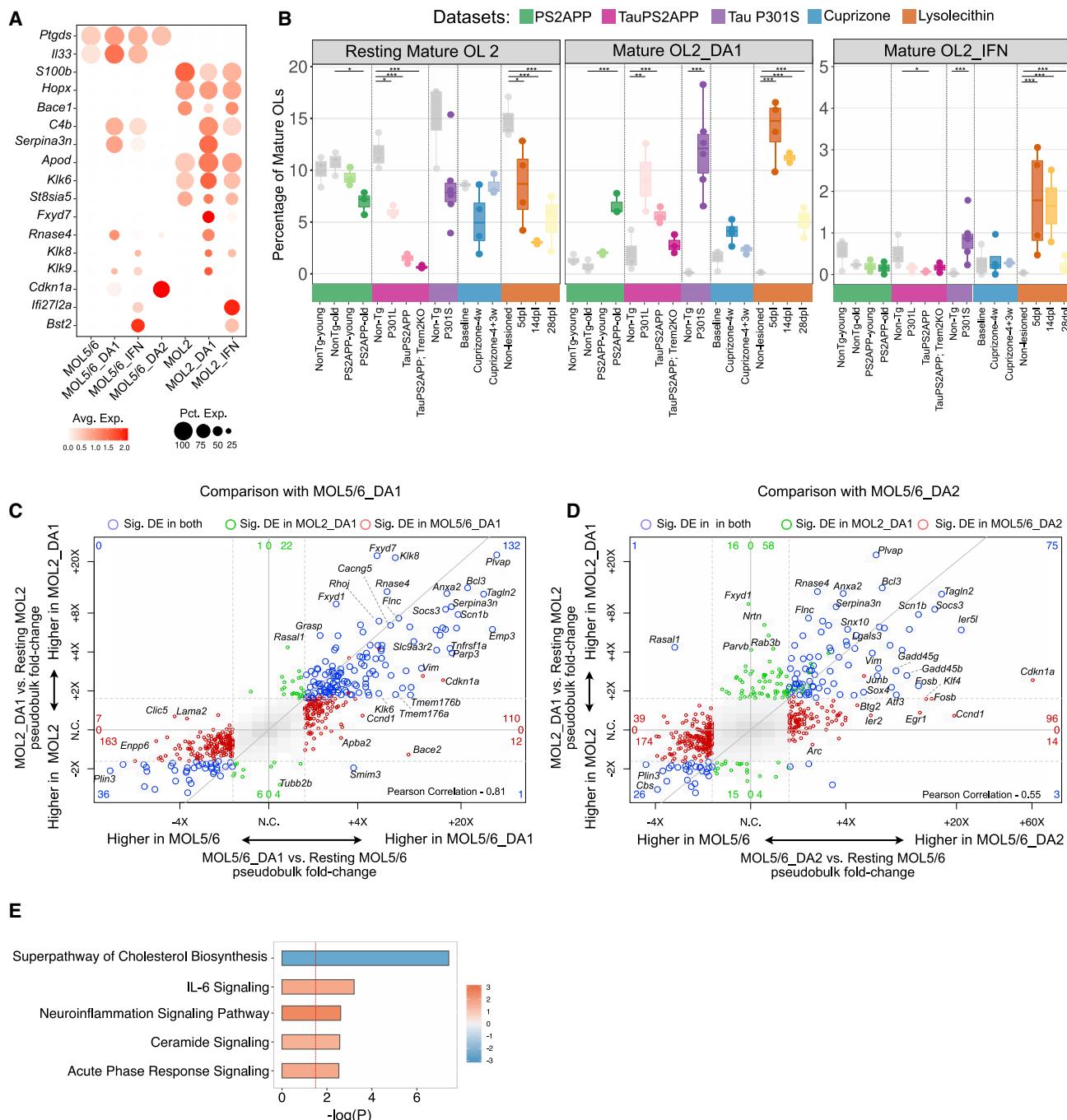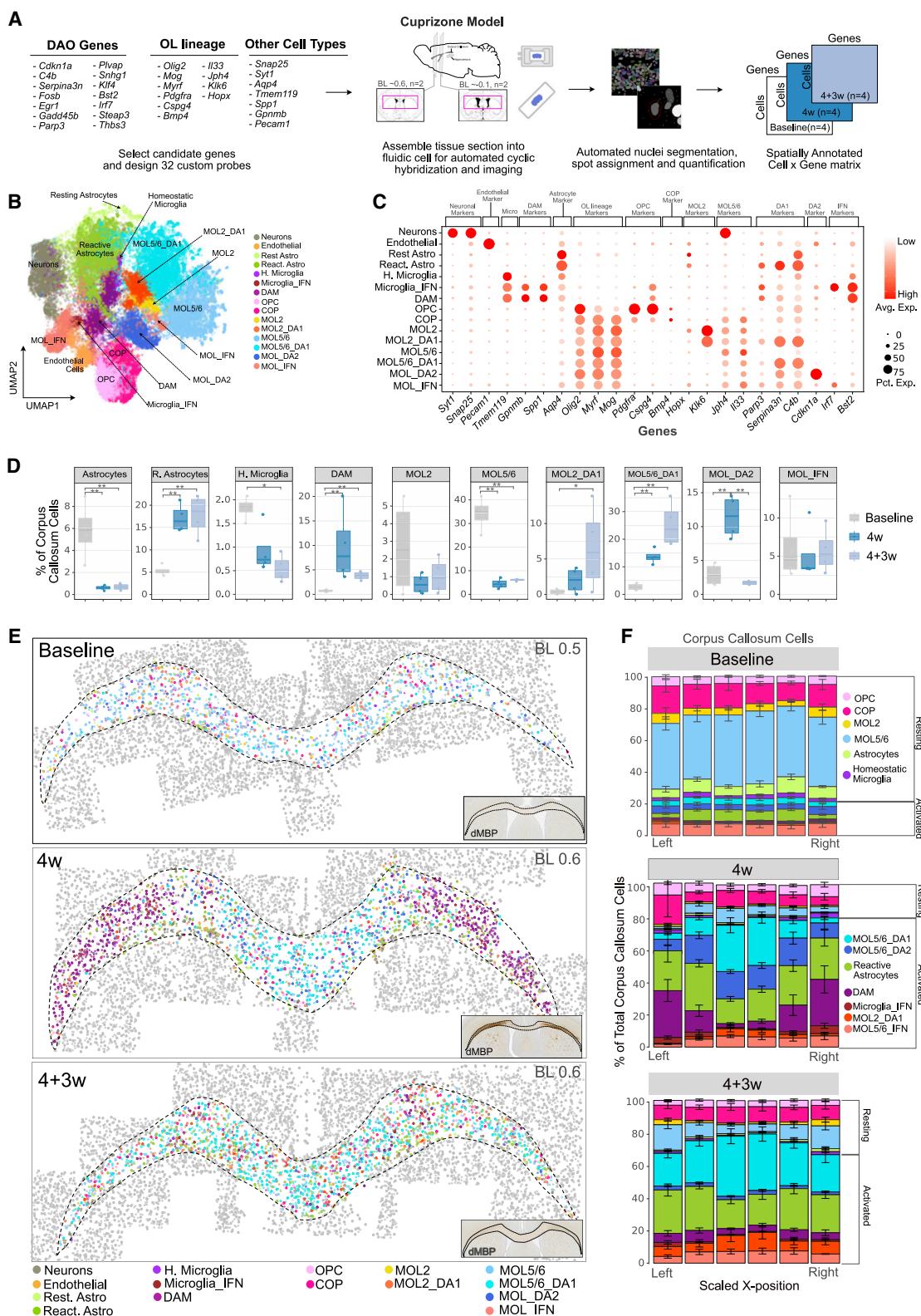
Figure 3. MOL2 demonstrates an inflammatory and interferon-related activation state that is shared with MOL5/6

(A) Dot plot of marker genes enriched in resting MOLs, MOL2_DA1, and MOL2_IFN.

(B) Relative proportions of resting MOL2, MOL2_DA1, and MOL2_IFN in each genotype or time point in the five separate datasets represented in the form of a boxplot. Differential abundance statistics: *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001 compared with baseline or non-transgenic condition. FDR, false discovery rate.

(C and D) Four-way comparison of DE genes in MOL2_DA1 (y axis) with DE genes in MOL5/6_DA1 (C; x axis) or MOL5/6_DA2 (D; x axis). Each point represents one gene colored by whether logFC ≥ 1.75 and FDR ≤ 0.05 in one or both differential expression analyses (red for MOL5/6_DA1 [C] or MOL5/6_DA2 [D], green for MOL2_DA1, or blue for both). Corresponding numbers of DE genes are shown near the borders of the plot. Diagonal line, y = x.

(E) Ingenuity pathway analysis of DE genes in MOL2_DA1. High Z scores depict predicted activation and low Z scores depict predicted inhibition of the respective pathway. All genes associated with the MOL2_DA1 state were included. See also Figures S4 and S5 and Table S2.

(legend on next page)

the corpus callosum between Bregma levels −0.1 and 0.6. The resulting expression matrix was filtered to remove low-quality cells and potential doublets (STAR Methods) (Figure S6A). No batch effects were observed in the dataset (Figure S6B).

To specifically validate our scRNA-seq dataset, which was generated from the corpus callosum, we computationally identified the boundary of the corpus callosum using the expression of neuronal and OL markers (Figure S6C; STAR Methods). Unsupervised clustering analysis of the resulting expression matrix identified all MOL and DAO subpopulations found in our single-cell meta-analysis: (1) MOL2 (*Klk6+*; *Hopx+*), (2) MOL5/6 (*Jph4*^{hi} and *Il33*^{hi}), (3) two *C4b*^{hi} *Serpina3n*^{hi} DA1 states corresponding to MOL5/6_DA1 and MOL2_DA1, (4) *Cdkn1a*^{hi} *C4b*^{lo} *Serpina3n*^{lo} DA2 population, and (5) *Irf7*^{hi} IFN state (Figures 4B and 4C). All clusters were composed of cells from all experimental batches (Figure S6B). At baseline, MOL5/6 made up for a larger proportion of OLs in the corpus callosum compared with MOL2. As expected, resting MOLs decrease and DAOs increase in proportion post cuprizone treatment (Figure 4D). A large majority of the remaining OLs at 4 weeks adopt a DA1 or DA2 phenotype. While the DA2 phenotype resolves at 4 + 3 weeks, a majority of the recovered OLs at 4 + 3 weeks still adopt a DA1 phenotype (Figure 4D). The persistence of the DA1 state was also observed in the scRNA-seq meta-analysis: the increase in proportion of the DA1 state at 4 + 3 weeks noted in the smFISH dataset may be a result of sensitivity differences between scRNA-seq and smFISH. Together, these results corroborate an alteration in the transcriptional profile of OLs during the demyelination and remyelination process.

In addition to DAOs, we also identified other clusters corresponding to homeostatic microglia (*Tmem119+*), DAMs (*Tmem119+*, *Gpnmb+*, *Spp1+*), astrocytes (*Aqp4+*), reactive astrocytes (*Aqp4+*, *Serpina3n+*, *C4b+*), and endothelial cells (*Pecam1+*) (Figure 4C). We replicated our previously reported increase in DAMs and reactive astrocytes at 4 weeks post cuprizone treatment compared with baseline (Figure 4D) (Shen et al., 2021). While microgliosis largely resolves at 4 + 3 weeks, astrogliosis persists into the recovery time point (Figure 4D).

Next, we investigated the spatial relationship of DAOs to the putative lesion site and to other demyelination-associated glial populations. Our data were collected from sections in which the stereotypic cuprizone-induced lesion appears at the lateral ends of the corpus callosum (Figure 4E, insets, and S6D). To quantify the spatial relationship of DAOs to this putative lesion site, we identified the midline of the corpus callosum and computed the distance of every cell to that midline (Figure S6C). To investigate how cell proportion changed with varying distances from this midline, we normalized this distance from the midline to go from −1 at the leftmost part to +1 at the rightmost part of the corpus callosum. At baseline, homeostatic MOLs, astrocytes, and microglia were all evenly distributed along the L-R axis (Figures 4E and 4F, top). At 4 weeks post cuprizone treatment, we observed abundant microglial proliferation and migration into the lesion site (Figures 4E and 4F, middle). Surviving or newly differentiated OLs at 4 weeks were situated away from the lesion site, to the midline of the corpus callosum, and had adopted either a DA1 or a DA2 phenotype. Interestingly, while DAMs were largely confined to the lesion site and DAOs away from the lesion site, reactive astrocytes were more evenly distributed along the L-R axis, outlining the boundary of the entire corpus callosum. At 4 + 3 weeks, as the microgliosis resolved, the lateral ends of the corpus callosum containing the lesion site were repopulated by DA1 OLs as noted by the largely even distribution of MOL5/6_DA1 and MOL2_DA1 along the L-R axis (Figures 4E and 4F, bottom). It is also noteworthy that the DA1 state persists into 4 + 3 weeks, whereas the DA2 reaches nearly complete resolution (Figures 4E, 4F, and S6E). Taken together, these results validate the DAO states observed in our scRNA-seq meta-analysis and put them in the spatial context of a cuprizone-induced lesion during demyelination and remyelination.

Cross-species integration of OL-lineage cells defines resting OL subtypes in human with some transcriptional similarity to mouse

Next, we wondered whether the disease-associated gene signatures identified in mouse models are also elevated in human disease. However, a necessary prerequisite to this analysis was to build an understanding of the heterogeneity of resting OLs in human and relate those to mouse subtypes. While there is a consensus in the literature on the heterogeneity of resting OL subtypes in mouse, a similar consensus for human OLs is lacking. To characterize baseline human OL heterogeneity and relate it to mouse subtypes, we performed an integrative analysis of control OLs derived from mouse and human datasets incorporated into our study.

To this end, we selected OL-lineage cells derived from control subjects in human AD (Cain et al., 2020; Grubman et al., 2019; Mathys et al., 2019) and MS (Absinta et al., 2021; Jäkel et al., 2019; Schirmer et al., 2019) datasets using canonical OL-lineage markers such as *SOX10*, *OLIG1*, *OLIG2*, *PDGFRA*, *VCAN*, *MOG*, *MYRF*, and *PLP1*. We similarly selected OL-lineage cells from the mouse datasets included in our study. Using 2,000 genes that are commonly variable across all datasets derived from both species, we integrated these datasets using Seurat's

---

Figure 4. *In situ* hybridization reveals distinct DA1 and DA2 subtypes in the cuprizone model

(A) Automated spatial RNA transcriptomics workflow used to validate candidate DAO gene expression and investigate spatial localization.

(B) UMAP plot highlighting the major glial subtypes recovered from the dataset across all experimental runs and time points. Cluster labels are represented in the form of a dendrogram built by performing hierarchical clustering on the average gene-expression profile for each cluster.

(C) Dot plot of marker genes enriched in DAO, DAM, and DAA subtypes in the corpus callosum.

(D) Relative proportions of DAO and other glial subtypes in the corpus callosum across the three time points. Differential abundance statistics: *FDR < 0.01, **FDR < 0.001.

(E) Spatial distribution of all major cell classes across a representative section for each time point in a single cuprizone experiment. Cells are colored by cell class as indicated (Bregma level 0.5 at baseline and 0.6 at 4 and 4 + 3 weeks). Insets mark the degraded myelin stain highlighting the putative lesion site.

(F) A stacked bar plot representing the proportions of cell types across ∼600-μm bins generated along the left-right axis of the corpus callosum. Bars and lines, mean ± SEM. See also Figure S6.
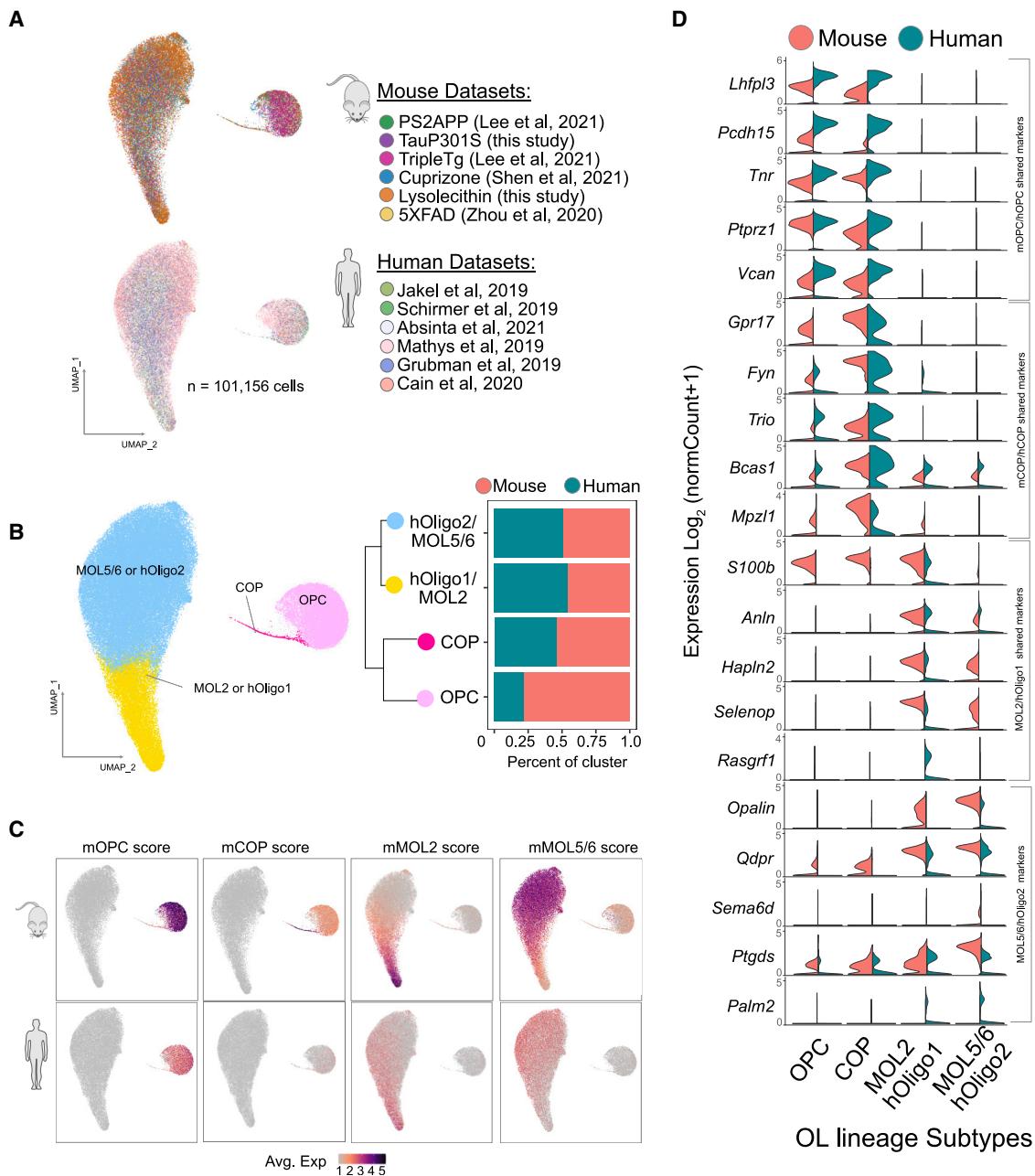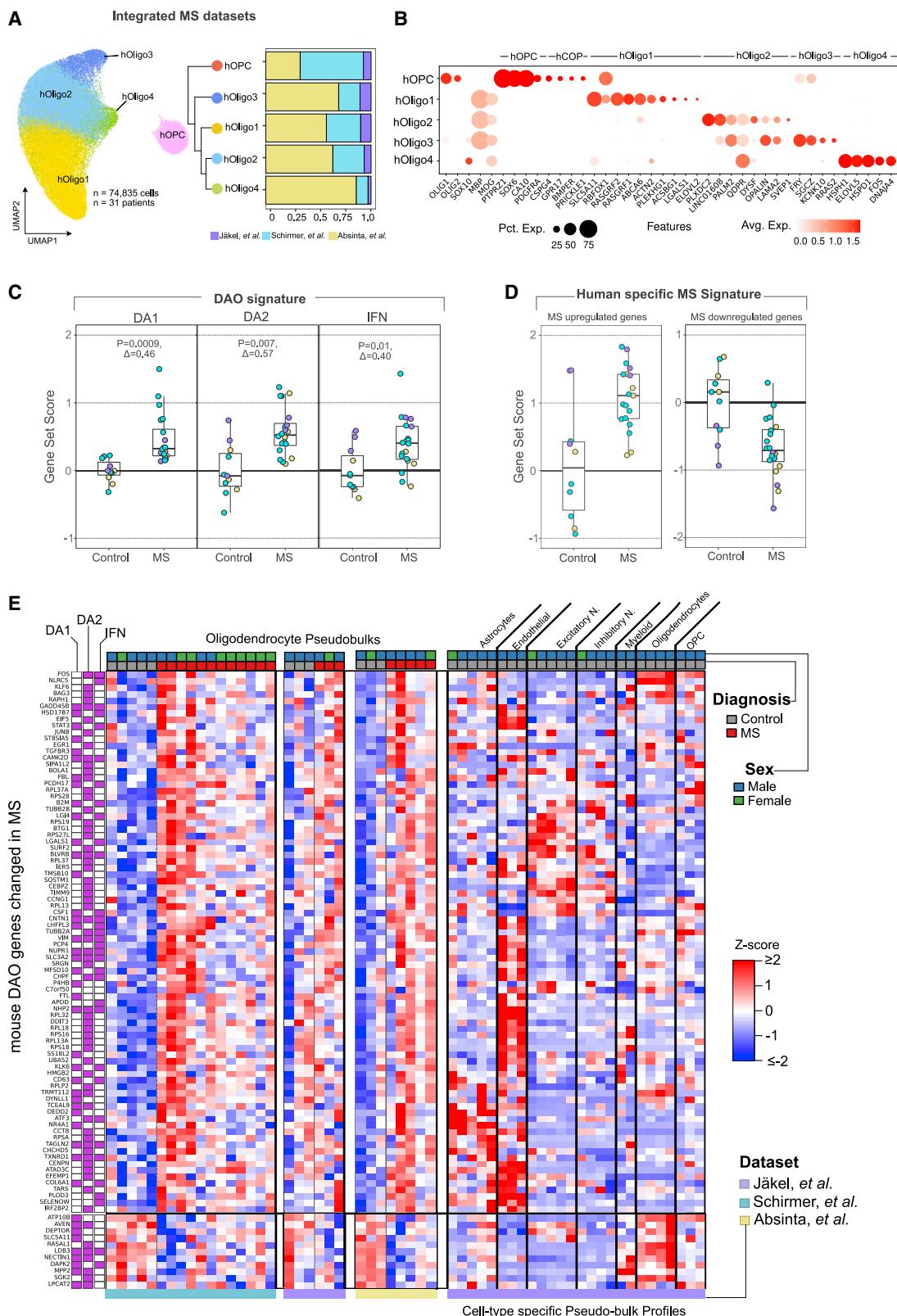
Figure 5. Cross-species integration of resting OL-lineage subtypes in human and mouse

(A) UMAP plot representing an integrated analysis of control OLs from six independent mouse and six independent human datasets colored by dataset and split by species.

(B) UMAP plot representing the clustering results from an integrated analysis of control OLs from human and mouse. Cluster labels are represented in the form of a dendrogram as in Figure 1C. The proportion of cells from each species for every cluster is represented in the bar plot on the right.

(C) Gene set score of mouse OL-lineage subtype markers represented in the form of a UMAP plot split by species.

(D) Split violin plots showing expression of conserved marker genes associated with resting OL-lineage subtypes in human and mouse. See also Figure S7.

MNN approach, resulting in a combined analysis of ~101,000 cells (Figure 5A). Subclustering of this integrated dataset resulted in four major clusters: OPC, COP, and two broad MOL subtypes (Figure 5B). Among the MOLs, we found that a subset of the human OLs co-clustered with MOL2, whereas others co-clustered with MOL5/6 (Figure S7A). We named these clusters hOligo1 and hOligo2, respectively. Similarly, using the gene set score of mouse MOL-specific gene sets, we found that MOL2-specific genes are enriched in hOligo1 and MOL5/6-specific gene sets are enriched in hOligo2 (Figures 5B and 5C). We

A Integrated MS datasets

B

C DAO signature

D Human specific MS Signature

E

(legend on next page)

highlight example conserved and unique genes between human and mouse in Figure 5D.

We used these clustering results to generate sample-specific pseudo-bulks of the human OL subtypes and performed a sub-type-specific differential expression analysis to find other markers that are enriched in hOligo1 and hOligo2 (Figure S7B). All subsequent annotations of human OL-lineage cells in MS and AD datasets were performed using these marker genes (Figure S7C).

Mouse DAO response is recapitulated in human OL responses in MS

OL-lineage cells are vulnerable in MS. To describe disease-specific transcriptional changes in MS OLs and understand their overlap with our analysis of mouse models, we performed an integrated analysis of OL-lineage cells from three published snRNA-seq datasets from MS patients (Absinta et al., 2021; Jä-kel et al., 2019; Schirmer et al., 2019). Integrated analyses revealed a cluster of OPCs and four MOLs (Figure 6A), including the two major resting human subtypes uncovered in the previous section (*RASGRF2* and *RASGRF1*+ hOligo1 and *PALM2*, *QDPR*, and *OPALIN*+ hOligo2) and two others that share markers with or branch out of hOligo2, which we dubbed hO-ligo3 (*FRY*, *RRAS2*+) and hOligo4 (*HSPH1*, *FOS*, *DNAJA4*+) (Figure 6B). Each OL cluster was represented in all datasets (Figure 6A, right, and Figure S8A) and was composed of cells from multiple patients except for hOligo4, a large proportion of which came from a single patient from Absinta et al. (2021) (Figure S8B). Each cluster was also identifiable in a separate independent analysis of each dataset except for hOligo4, which clustered out only in Schirmer et al. (2019) and Absinta et al. (2021) (Figure S8C). Markers of MOL subtypes were robustly expressed across all datasets (Figure S8D). While no differences in abundance of hOligo1 and hOligo2 were observed between MS and control patients, both hOligo3 and hOligo4 showed modest elevation in proportion in MS patients (Figure S8E). However, these differences were significant only in Schirmer et al. (2019) for hOligo3 and Absinta et al. (2021) for hOligo4.

To assess whether the gene sets associated with the DA1, DA2, and IFN OLs identified in mouse models were also elevated in OLs from MS patients, we aggregated all cells from OL subtypes hOligo1–4 into a single pseudo-bulk expression profile (STAR Methods) and scored the pseudo-bulk profiles for expression of the mouse DAO genes associated with DA1, DA2, and IFN states (Figure 6C). We detected an increased expression of DA1-, DA2-, and IFN-related gene set scores in OLs obtained

from MS patients across all datasets, although the differences were largest in Absinta et al. (2021) and Schirmer et al. (2019) (Figure 6C).

We next sought to assess (1) which mouse DAO genes are conserved in a separate and independent differential expression analysis of human OLs in MS and (2) if there are additional genes that reveal human-specific OL activation in MS. To this end, we generated pseudo-bulk RNA-seq profiles by separately aggregating single cells from control and MS patients in hOligo1 and hOligo2. We aggregated cells belonging to hOligo3 and hOligo4 together with hOligo2 as they largely shared markers and seemed to be disease-elevated states of hOligo2 (Figure S8E). We then performed differential gene expression analysis between control and MS samples within both hOligo1 and hOligo2 separately and mapped their expression back to UMAP coordinates (Figure S8F). We found the largest number of DE genes between control and MS patients in the *OPALIN*+ hOligo2 (Figures S9A and S6D), which were also shared with the MS versus control DE profile in *RASGRF1*+ hOligo1 (Figure S9B). These genes were elevated across all three datasets (Figure 6D). A number of these DE genes are shared with mouse DA1, DA2, and IFN states (Figure 6E, denoted in purple, and Figures S9C–S9E). Of the shared genes, we found DA2 genes related to the EIF2 signaling pathway, such as *EIF5*, *ATF3*, *EIF1*, *DDIT3*, and *RPL13*, and to the unfolded protein response, such as *SQSTM1* and *CEBPZ* (Figures 6E and S9D). Similarly, consistent with the upregulation of antigen presentation pathway genes in DA1 and IFN OLs from mouse models, OLs derived from MS patients also upregulated genes such as *NLRC5*, *B2M*, and *PSMB6* (Figures 6E and S9A). This is in line with several reports of antigen presenting capabilities of OLs and precursors (Falcão et al., 2018; Kirby and Castelo-Branco, 2020; Kirby et al., 2019).

MS OLs also upregulated unique genes hinting at human-specific OL activation (Figures S9C–S9E). Among those were genes indicative of cellular stress, such as the unfolded protein response (*PFDN4*, *UBB*, *UBC*, *UBA1*) and heat-shock response (*HSPA4L*, *HSPA4*, *HSPA5*, *HSPH1*, *ST3*, *DNAJA1*), and genes indicative of metabolic shift, such as oxidative phosphorylation (*COX5B*, *COX7B*) and the sirtuin pathway (*SIRT2*, *SOD1*, *SLC25A6*). In addition to the antigen presentation pathway genes shared with mouse models, MS OLs also upregulated genes involved in phagosome maturation, such as *PRDX5*, *VAMP3*, and *ATP6AP1*. Conversely, we observed downregulation of cell adhesion molecules such as *NRXN3*, *CDH1*, *CDH23*, *CDH11*, and *NLGN3*; ECM proteins such as

Figure 6. Integration of human OLs from MS datasets reveals that select orthologs of mouse DAO genes are elevated in MS
(A) UMAP representation of the integrated human OLs from three published MS datasets. Cluster labels are represented in the form of a dendrogram as in Figure 1C. Percentage of cells from each dataset contributing to each cluster is highlighted in the bar plot.
(B) Dot plot of marker genes enriched in human OL-lineage subtypes across the integrated MS datasets.
(C) Distribution of scores for mouse-derived gene sets across human OLs derived from MS and control individuals. Gene set scores were centered to control group. P represents p (unadjusted), t test.
(D) DE gene set scores of genes derived from an integrated differential expression analysis comparing pseudo-bulks of hOligo2 from MS patients with those from control individuals.
(E) Heatmap of the mouse DAO genes also differentially expressed in hOligo2 across control and MS patients. The full list of DE genes is in Figure S9. Columns on the left are colored in purple by whether the gene was classified as a DA1, DA2 ,or IFN gene in the mouse-specific analysis. See also Figures S8 and S9 and Table S3.
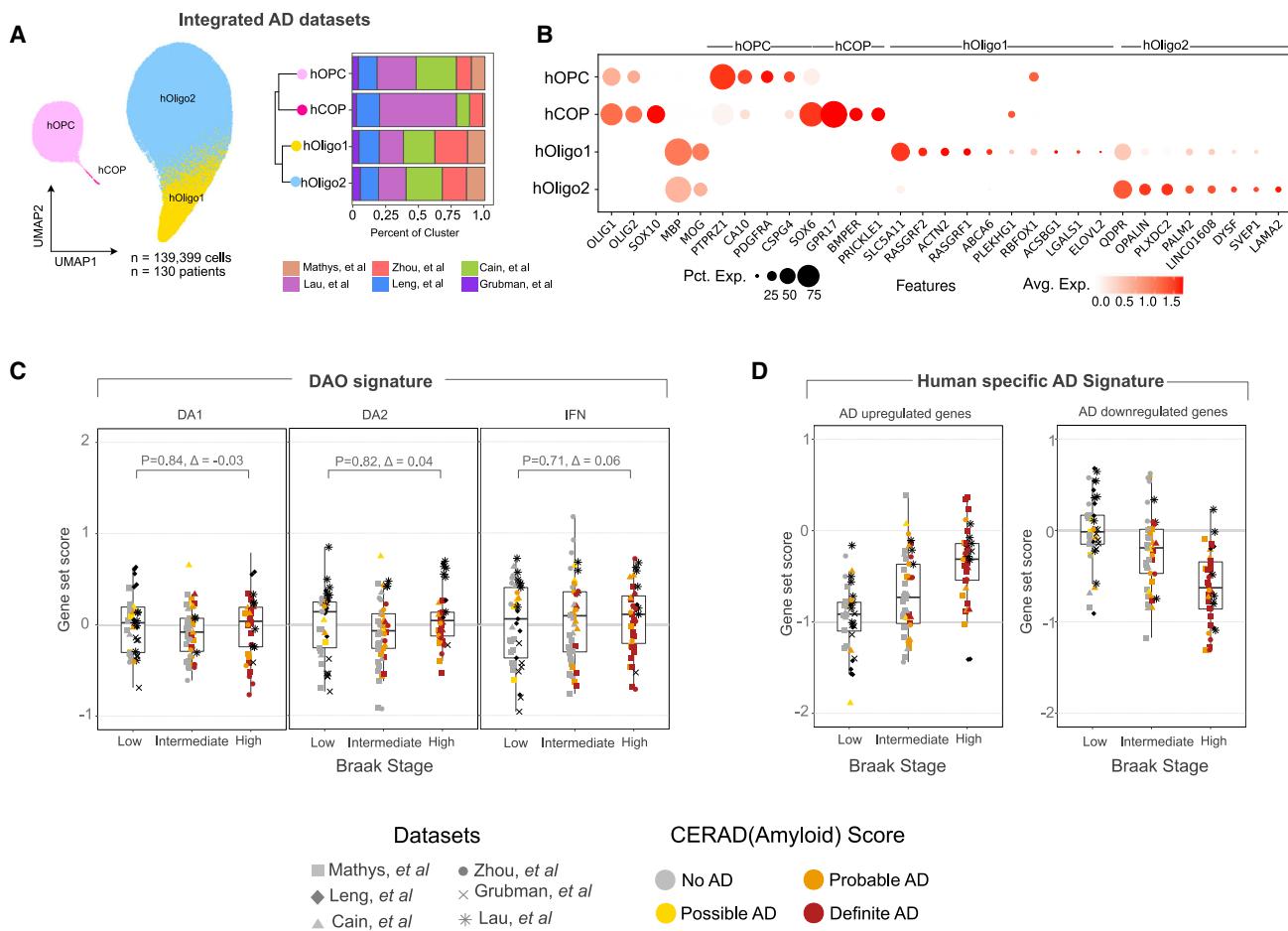
Figure 7. Integration of six AD datasets reveals that human OLs exhibit an AD-associated DE profile that is largely distinct from the mouse DAO signature

(A) UMAP representation of the integrated human OLs across six AD datasets. Cluster labels are represented in the form of a dendrogram as in Figure 1C. Percentage of cells from each dataset contributing to each cluster is highlighted in the bar plot.

(B) Dot plot of marker genes enriched in human OL-lineage subtypes across the integrated AD datasets.

(C) Distribution of scores for mouse-derived gene sets across hOligo2 undergoes little to no change in high-pathology AD patients compared with low-pathology individuals. Gene set scores were centered to the low-pathology group. Δ represents the difference between high pathology and low pathology. P represents p (unadjusted), t test.

(D) DE gene set scores of genes derived from a differential expression analysis comparing pseudo-bulks of hOligo2 from high-pathology AD patients with those from low-pathology individuals. Only upregulated genes are shown on the left and only downregulated genes are shown on the right. All gene set scores were centered to the low-pathology group. A heatmap of all DE genes is in Figure S10. See also Figure S10 and Table S3.

ADAMTS14, *ADAMTS17*, and *COL16A1*; specific contactins required for the organization of axonal paranodes, such as *CNTNAP2*; and genes required for formation of the nodes of Ranvier, such as *GLDN* (Çolakoğlu et al., 2014; Feinberg et al., 2010; Maluenda et al., 2016) (Figure S9A). Collectively, these results suggest that OLs in MS upregulate aspects of the mouse DAO gene signature corresponding to DA1, DA2, and IFN OLs, while also upregulating human-specific genes. Although our data do not indicate that these activations occurred in distinct cell populations in MS patient-derived OLs (Figure S8F), this may be a result of the sample-level integration performed in our analysis to mitigate patient-specific clustering.

Human AD OL response is largely distinct

Extensive evidence from multiple studies has reported the vulnerability of OLs in AD. However, it is not clear how these responses relate to one another and which aspects of these responses might be recapitulated in animal models. We performed an integrated analysis of OL-lineage cells from a number of recently published snRNA-seq datasets from AD patients to (1) understand if the responses observed in our mouse models are observed in human AD and (2) identify a shared AD-associated OL gene signature across multiple studies. To this end, we identified human OL-lineage cells from six snRNA-seq datasets derived from AD patients (Cain et al., 2020; Grubman et al., 2019; Lau et al., 2020; Leng et al., 2021; Mathys et al., 2019;

Zhou et al., 2020) and integrated ~139,000 cells together using Seurat's MNN approach (Figure S10A).

Interestingly, while these datasets were collected from different brain regions, including the dorsolateral prefrontal cortex, superior frontal gyrus, hippocampus, and entorhinal cortex, we observed the same subtypes of MOLs in the integrated analysis: a *RASGRF2* and *RASGRF1*+ hOligo1 and a *PLXDC2*, *PALM2*, and *OPALIN*+ hOligo2 (Figures 7A and 7B). Each cluster of OLs was represented in all datasets (Figure 7A, right, and Figure S10A) and was also recovered in a separate independent analysis of each AD dataset, suggesting that the integration results were not dominated by signals from a single or a few datasets (Figure S10B). Each cluster was composed of cells from a majority of the 31 patients (cognitively normal and AD patients) incorporated into the analysis (Figure S10C). MOL markers used to define clusters were robustly expressed across all datasets (Figure S10D). We did not observe a state similar to hOligo3 or hOligo4 in the AD datasets, which may be a result of differences in brain regions profiled or pathology observed in MS and AD.

We next sought to assess whether the gene sets associated with DA1, DA2, and IFN OLs identified in our mouse meta-analysis are also elevated in OLs derived from AD patients. Thus, we stratified AD patients into low (Braak stage in 0, I, and II and CERAD/plaque score in no AD or possible AD), intermediate (Braak stages III and IV and CERAD/plaque score probable AD), and high (Braak stages V and VI and CERAD/plaque score in probable AD and definite AD) pathology groups and aggregated all the MOLs into a single pseudo-bulk expression profile (STAR Methods). Scoring the pseudo-bulk profiles for expression of the mouse DAO genes associated with DA1, DA2, and IFN states revealed no differences in gene set scores among low-, intermediate-, and high-pathology individuals (Figure 7C).

Given this lack of correspondence between mouse DAO and human AD response, we sought to define an AD-associated OL signature that is shared across all the datasets integrated in our analysis. To this end, we aggregated single cells in each broad MOL cluster hOligo1 and hOligo2 to generate pseudo-bulk RNA-seq profiles across the three patient groups defined above. For both MOL subtypes, genome-wide differential expression analysis was performed between cells derived from high-pathology individuals and low-pathology individuals. Within hOligo1, we identified six DE genes between high- and low-pathology individuals. Performing the same analysis with hOligo2 yielded 41 genes increased and 43 genes decreased in hOligo2 derived from high-pathology individuals compared with those derived from low-pathology individuals (Figure S10E). These genes were largely distinct from the mouse-derived DAO signature associated with DA1 and DA2 OLs. However, the gene sets derived from the human meta-analysis are robustly changed in OLs derived from high-pathology AD individuals across all datasets (Figure 7D). Although hOligo1 yielded fewer DE genes between low- and high-pathology individuals, a similar trend in fold changes was observed in hOligo1 (Figure S10F). Observing the average expression of this AD-elevated gene set along the UMAP coordinates highlights the overall enrichment of this AD-associated signature in both hOligo1 and hOligo2 (Figure S10G).

Among the genes elevated in OLs in AD are lipid biosynthesis and metabolism genes such as *MID1IP1* and *ABCA6*, cytoskeletal proteins such as *ANLN*, and proteins involved in cell adhesion such as *CTNNA1*, *RELN*, and *CLDN15*. Only five genes were shared between the human AD-associated gene signature and the mouse model-derived DAO signature. These included cell division genes such as *PNPT1*, *SSB*, and *MITD1* and cell-cycle regulators *RGCC* and *CPQ*. Taken together, these results provide a set of commonly activated genes in OLs derived from AD patients and highlight their distinct transcriptional response from that observed in MS patients and mouse models of both diseases.

DISCUSSION

Here, we performed a multi-dataset integration of OL-lineage cells obtained from diverse datasets representing mouse models of neurodegenerative disease and *postmortem* data from patients. Our study highlights shared DAO states and gene expression programs across mouse models, places these transcriptomic states in an anatomical context *in situ*, and establishes the relevance of the findings in human disease.

Among the DAOs identified by our analysis, DA1 OLs were universally induced across all analyzed models. These OLs upregulated expression of immunomodulatory molecules, including cytokines, CD receptors, MHC class I and II molecules, and complement. Similar OLs have been observed in the experimental autoimmune encephalomyelitis (EAE) model and in human MS lesions and have been hypothesized to increase OL susceptibility to cytotoxic CD8+ T cells (Falcão et al., 2018; Kirby and Castelo-Branco, 2020; Kirby et al., 2019). In addition, DA1 OLs also upregulate *Tnfrsf1a*, which harbors a number of MS susceptibility alleles (International MS Genetics Consortium et al., 2009). These findings, in the context of previously published studies, suggest that this immunogenic state may be damaging in disease, providing further evidence that OLs may not only be targets of the immune system but may also have immunomodulatory properties that influence disease outcomes.

Furthermore, we show that this immunogenic phenotype is also observed in amyloid and tau models of AD and may represent a common neuroinflammatory mechanism across diverse neurodegenerative states. Similar *C4b*+ and *Serpina3n*+ OLs were also described in the 5XFAD model (Zhou et al., 2020). This immune phenotype may result from pathological aggregates of amyloid or tau, increase in inflammatory cytokines, or axonal damage/degeneration signals that are induced in most of the analyzed models. The strong activation of the DA1 signature in the TauPS2APP mice, which contain abundant dystrophic neurites (Lee et al., 2021b) compared with the PS2APP or TauP301L mice, suggests that axonal damage may strongly induce this phenotype in OLs wrapping dystrophic axons.

In stark contrast to DA1 OLs, DA2 OLs upregulate pathways that influence OL survival. The upregulation of EIF2 signaling in the DA2 OLs is noteworthy. Previous studies have identified loss-of-function mutations in EIF2B in patients with vanishing white matter disease (VWMD) (Leegwater et al., 2001; Li et al., 2004). Furthermore, specific impairment of EIF2B activity in OLs has been shown to reproduce the pathology of VWMD in

mice (Lin et al., 2014). Other pathways upregulated by DA2 OLs, such as mTOR and CNTF signaling, are known to promote survival and differentiation of OPCs as well as myelination, hinting at a possible cytoprotective effect of DA2 OLs (Grier et al., 2017; Lebrun-Julien et al., 2014; Louis et al., 1993; Sanz-Rodriguez et al., 2018; Stankoff et al., 2002; Tyler et al., 2009). However, the expression of cell-cycle arrest and pro-apoptotic markers such as *Cdkn1a*, *Gadd45a*, and *Bax* also hints at an ongoing struggle for survival.

Notably, all DAO states are accompanied by a downregulation in the components of the cholesterol biosynthetic pathway. Cholesterol biosynthesis supports myelin biogenesis and axon ensheathment (Berghoff et al., 2017; Saher et al., 2005). Yet, recent evidence suggests that the inhibition of a subset of enzymes within the cholesterol biosynthesis pathway promotes OL formation via the accumulation of selected sterols (Hubler et al., 2018). Strikingly, *Ebp*, one such enzyme, is specifically downregulated in the DA2 OLs.

Spatial analysis of DA1 and DA2 OLs in the cuprizone model highlighted striking differences in the spatial patterns of DA1 and DA2 states during remyelination. While both DAOs are established outside the microglia-infiltrated lesion site during demyelination, DA1 OLs persist into the 4 + 3 week recovery time point, spreading into the lesion site, following the resolution of microgliosis (Figure 4E). This persistence, which was also observed in our single-cell meta-analyses, suggests that, despite substantial recovery of OL numbers during remyelination as measured by electron microscopy (EM) (Shen et al., 2021), the newly generated OLs do not default to a resting transcriptional state. This may not only have implications for neuroregeneration and repair but also provide biomarkers for early remyelination. Whether these new OLs still expressing disease-associated genes persist much longer after 4 + 3 weeks or lose their ability to remyelinate over repeated rounds of de- and remyelination remains an open question. Further studies with longer endpoints will be required to understand the specific functional roles of different MOLs in disease progression and/or repair.

Our analysis of human snRNA-seq studies from MS highlighted both similarities and differences with the analyzed remyelination mouse models. While these models do not have a strong adaptive immune system component present in MS, orthologs of DA1 and DA2 genes were elevated in MS OLs, highlighting both DA1- and DA2-like activation in MS. However, some orthologs of genes elevated in MS were not altered in the models, highlighting features of human-specific activation. Among the human-specific DE gene pathways in MS OLs were those indicative of cellular stress and inflammatory pathology, reflecting aspects of MS that are not captured in remyelination models. However, we also identified upregulation of putatively protective pathways, such as sirtuin signaling genes that are known to have a role in OL differentiation and remyelination (Beirowski et al., 2011; Ma et al., 2022). The differential expression profile we observed in AD was almost entirely distinct from the DAO profile observed in mouse models, except for a handful of genes. This is similar to human AD microglia (Srinivasan et al., 2020), which were found to have a signature largely distinct from the DAM signature identified in mouse models (Deczkowska et al., 2018; Friedman et al., 2018; Keren-Shaul et al., 2017). Collectively, our multi-dataset integration study uncovered

three major DAO states and established their relevance in human disease. More broadly, these gene expression data and analyses will suggest therapeutic approaches for remyelination strategies aimed at modifying disease outcomes by dampening negative states and enhancing beneficial ones.

Limitations of the study

While the differences between mouse model-related DAO signature and patient-related AD signature could be a result of intrinsic differences in mouse and human OL responses, the activation of many of the DAO genes in MS brains suggests that this may not be the only reason. An alternative possibility is that the AD-related OL pathology may be more pronounced and similar to mouse models in white matter areas, which remain unrepresented in the published datasets analyzed in our study. Additional profiles with increased resolution across various stages of AD and across many brain regions may shed further light on AD-related OL changes. Our catalog of DAOs and their associated gene programs will facilitate such future studies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  ○ Lead contact
  ○ Materials availability
  ○ Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  ○ Animals
- METHOD DETAILS
  ○ Preparation of single-cell suspensions from the lysolecithin treated animals
  ○ Preparation of single-cell suspensions from the TauP301S animals
  ○ Single-cell RNA-seq library preparation and sequencing
  ○ *In situ* hybridization, imaging and image analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  ○ Raw data analysis
  ○ Mouse scRNA-seq meta-analysis
  ○ Pseudo-bulk analysis of sc/snRNA-seq datasets
  ○ Differential expression (DE) analysis from pseudobulk sample
  ○ Gene set analysis
  ○ Human scRNA-seq meta-analysis
  ○ Differential expression (DE) analysis from pseudobulk sample to obtain an MS and AD-associated oligodendrocyte activation signature
  ○ Analysis of ISH data

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.111189.

REFERENCES

Absinta, M., Maric, D., Gharagozloo, M., Garton, T., Smith, M.D., Jin, J., Fitzgerald, K.C., Song, A., Liu, P., Lin, J.-P., et al. (2021). A lymphocyte–microglia–astrocyte axis in chronic active multiple sclerosis. Nature *597*, 709–714. https://doi.org/10.1038/s41586-021-03892-7.

Bartzokis, G. (2011). Alzheimer's disease as homeostatic responses to age-related myelin breakdown. Neurobiol. Aging *32*, 1341–1371. https://doi.org/10.1016/j.neurobiolaging.2009.08.007.

Beirowski, B., Gustin, J., Armour, S.M., Yamamoto, H., Viader, A., North, B.J., Michán, S., Baloh, R.H., Golden, J.P., Schmidt, R.E., et al. (2011). Sir-two-homolog 2 (Sirt2) modulates peripheral myelination through polarity protein Par-3/atypical protein kinase C (aPKC) signaling. Proc. Natl. Acad. Sci. USA *108*, E952–E961. https://doi.org/10.1073/pnas.1104969108.

Berghoff, S.A., Gerndt, N., Winchenbach, J., Stumpf, S.K., Hosang, L., Odoardi, F., Ruhwedel, T., Böhler, C., Barrette, B., Stassart, R., et al. (2017). Dietary cholesterol promotes repair of demyelinated lesions in the adult brain. Nat. Commun. *8*, 14241. https://doi.org/10.1038/ncomms14241.

Braak, H., and Braak, E. (1996). Development of Alzheimer-related neurofibrillary changes in the neocortex inversely recapitulates cortical myelogenesis. Acta Neuropathol. *92*, 197–201. https://doi.org/10.1007/s004010050508.

Cain, A., Taga, M., McCabe, C., Hekselman, I., White, C.C., Green, G., Rozenblatt-Rosen, O., Zhang, F., Yeger-Lotem, E., Bennett, D.A., et al. (2020). Multicellular communities are perturbed in the aging human brain and with Alzheimer's disease. Preprint at bioRxiv. https://doi.org/10.1101/2020.12.22.424084.

Caso, F., Agosta, F., Mattavelli, D., Migliaccio, R., Canu, E., Magnani, G., Marcone, A., Copetti, M., Falautano, M., Comi, G., et al. (2015). White matter degeneration in atypical alzheimer disease. Radiology *277*, 162–172. https://doi.org/10.1148/radiol.2015142766.

Çolakoğlu, G., Bergstrom-Tyrberg, U., Berglund, E.O., and Ranscht, B. (2014). Contactin-1 regulates myelination and nodal/paranodal domain organization in the central nervous system. Proc. Natl. Acad. Sci. USA *111*, E394–E403. https://doi.org/10.1073/pnas.1313769110.

International MS Genetics Consortium; Jia, X., Wang, J., de Bakker, P.I.W., Ottoboni, L., Aggarwal, N.T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C., et al. (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat. Genet. *41*, 776–782. https://doi.org/10.1038/ng.401.

Crawford, A.H., Tripathi, R.B., Richardson, W.D., and Franklin, R.J.M. (2016). Developmental origin of oligodendrocyte lineage cells determines response to

demyelination and susceptibility to age-associated functional decline. Cell Rep. *15*, 761–773. https://doi.org/10.1016/j.celrep.2016.03.069.

Deczkowska, A., Keren-Shaul, H., Weiner, A., Colonna, M., Schwartz, M., and Amit, I. (2018). Disease-associated microglia: a universal immune sensor of neurodegeneration. Cell *173*, 1073–1081. https://doi.org/10.1016/j.cell.2018.05.003.

Falcão, A.M., van Bruggen, D., Marques, S., Meijer, M., Jäkel, S., Agirre, E., Samudyata, Floriddia, E.M., Floriddia, E.M., Vanichkina, D.P., et al. (2018). Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. Nat. Med. *24*, 1837–1844. https://doi.org/10.1038/s41591-018-0236-y.

Fazeli, E., Roy, N.H., Follain, G., Laine, R.F., von Chamier, L., Hänninen, P.E., Eriksson, J.E., Tinevez, J.-Y., and Jacquemet, G. (2020). Automated cell tracking using StarDist and TrackMate. F1000Res. *9*, 1279. https://doi.org/10.12688/f1000research.27019.2.

Feinberg, K., Eshed-Eisenbach, Y., Frechter, S., Amor, V., Salomon, D., Sabanay, H., Dupree, J.L., Grumet, M., Brophy, P.J., Shrager, P., and Peles, E. (2010). A glial signal consisting of gliomedin and NrCAM clusters axonal Na+ channels during the formation of nodes of ranvier. Neuron *65*, 490–502. https://doi.org/10.1016/j.neuron.2010.02.004.

Floriddia, E.M., Lourenço, T., Zhang, S., van Bruggen, D., Hilscher, M.M., Kukanja, P., Gonçalves Dos Santos, J.P., Altınkök, M., Yokota, Y., Llorens-Bobadilla, E., et al. (2020). Distinct oligodendrocyte populations have spatial preference and different responses to spinal cord injury. Nat. Commun. *11*, 5860. https://doi.org/10.1038/s41467-020-19453-x.

Friedman, B.A., Srinivasan, K., Ayalon, G., Meilandt, W.J., Lin, H., Huntley, M.A., Cao, Y., Lee, S.-H., Haddick, P.C.G., Ngu, H., et al. (2018). Diverse brain myeloid expression profiles reveal distinct microglial activation states and aspects of Alzheimer's disease not evident in mouse models. Cell Rep. *22*, 832–847. https://doi.org/10.1016/j.celrep.2017.12.066.

Fünfschilling, U., Supplie, L.M., Mahad, D., Boretius, S., Saab, A.S., Edgar, J., Brinkmann, B.G., Kassmann, C.M., Tzvetanova, I.D., Möbius, W., et al. (2012). Glycolytic oligodendrocytes maintain myelin and long-term axonal integrity. Nature *485*, 517–521. https://doi.org/10.1038/nature11007.

Gensert, J.M., and Goldman, J.E. (1997). Endogenous progenitors remyelinate demyelinated axons in the adult CNS. Neuron *19*, 197–203. https://doi.org/10.1016/s0896-6273(00)80359-1.

Götz, J., Chen, F., Barmettler, R., and Nitsch, R.M. (2001). Tau filament formation in transgenic mice expressing P301L tau. J. Biol. Chem. *276*, 529–534. https://doi.org/10.1074/jbc.m006531200.

Grier, M.D., West, K.L., Kelm, N.D., Fu, C., Does, M.D., Parker, B., McBrier, E., Lagrange, A.H., Ess, K.C., and Carson, R.P. (2017). Loss of mTORC2 signaling in oligodendrocyte precursor cells delays myelination. PLoS One *12*, e0188417. https://doi.org/10.1371/journal.pone.0188417.

Griffiths, I., Klugmann, M., Anderson, T., Yool, D., Thomson, C., Schwab, M.H., Schneider, A., Zimmermann, F., McCulloch, M., Nadon, N., and Nave, K.A. (1998). Axonal swellings and degeneration in mice lacking the major proteolipid of myelin. Science *280*, 1610–1613. https://doi.org/10.1126/science.280.5369.1610.

Grubman, A., Chew, G., Ouyang, J.F., Sun, G., Choo, X.Y., McLean, C., Simmons, R.K., Buckberry, S., Vargas-Landin, D.B., Poppe, D., et al. (2019). A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. Nat. Neurosci. *22*, 2087–2097. https://doi.org/10.1038/s41593-019-0539-4.

Grueninger, F., Bohrmann, B., Czech, C., Ballard, T.M., Frey, J.R., Weidensteiner, C., von Kienlin, M., and Ozmen, L. (2010). Phosphorylation of Tau at S422 is enhanced by Aβ in TauPS2APP triple transgenic mice. Neurobiol. Dis. *37*, 294–306. https://doi.org/10.1016/j.nbd.2009.09.004.

Haimon, Z., Volaski, A., Orthgiess, J., Boura-Halfon, S., Varol, D., Shemer, A., Yona, S., Zuckerman, B., David, E., Chappell-Maor, L., et al. (2018). Re-evaluating microglia expression profiles using RiboTag and cell isolation strategies. Nat. Immunol. *19*, 636–644. https://doi.org/10.1038/s41590-018-0110-6.

Hubler, Z., Allimuthu, D., Bederman, I., Elitt, M.S., Madhavan, M., Allan, K.C., Shick, H.E., Garrison, E., T Karl, M., Factor, D.C., et al. (2018). Accumulation of

8, 9-unsaturated sterols drives oligodendrocyte formation and remyelination. Nature *560*, 372–376. https://doi.org/10.1038/s41586-018-0360-3.

Ishii, A., Furusho, M., Macklin, W., and Bansal, R. (2019). Independent and cooperative roles of the Mek/ERK1/2-MAPK and PI3K/Akt/mTOR pathways during developmental myelination and in adulthood. Glia *67*, 1277–1295. https://doi.org/10.1002/glia.23602.

Jäkel, S., Agirre, E., Mendanha Falcão, A., van Bruggen, D., Lee, K.W., Knuesel, I., Malhotra, D., Ffrench-Constant, C., Williams, A., and Castelo-Branco, G. (2019). Altered human oligodendrocyte heterogeneity in multiple sclerosis. Nature *566*, 543–547. https://doi.org/10.1038/s41586-019-0903-2.

Keren-Shaul, H., Spinrad, A., Weiner, A., Matcovitch-Natan, O., Dvir-Szternfeld, R., Ulland, T.K., David, E., Baruch, K., Lara-Astaiso, D., Toth, B., et al. (2017). A unique microglia type Associated with restricting development of Alzheimer's disease. Cell *169*, 1276–1290.e17. https://doi.org/10.1016/j.cell.2017.05.018.

Kirby, L., and Castelo-Branco, G. (2020). Crossing boundaries: interplay between the immune system and oligodendrocyte lineage cells. Semin. Cell Dev. Biol. *116*, 45–52. https://doi.org/10.1016/j.semcdb.2020.10.013.

Kirby, L., Jin, J., Cardona, J.G., Smith, M.D., Martin, K.A., Wang, J., Strasburger, H., Herbst, L., Alexis, M., Karnell, J., et al. (2019). Oligodendrocyte precursor cells present antigen and are cytotoxic targets in inflammatory demyelination. Nat. Commun. *10*, 3887. https://doi.org/10.1038/s41467-019-11638-3.

van der Knaap, M.S., and Bugiani, M. (2017). Leukodystrophies: a proposed classification system based on pathological changes and pathogenetic mechanisms. Acta Neuropathol. *134*, 351–382. https://doi.org/10.1007/s00401-017-1739-1.

Lau, S.-F., Cao, H., Fu, A.K.Y., and Ip, N.Y. (2020). Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. Proc. Natl. Acad. Sci. USA *117*, 25800–25809. https://doi.org/10.1073/pnas.2008762117.

Lebrun-Julien, F., Bachmann, L., Norrmén, C., Trötzmüller, M., Köfeler, H., Rüegg, M.A., Hall, M.N., and Suter, U. (2014). Balanced mTORC1 activity in oligodendrocytes is required for accurate CNS myelination. J. Neurosci. *34*, 8432–8448. https://doi.org/10.1523/jneurosci.1105-14.2014.

Lee, S.-H., Meilandt, W.J., Xie, L., Gandham, V.D., Ngu, H., Barck, K.H., Rezzonico, M.G., Imperio, J., Lalehzadeh, G., Huntley, M.A., et al. (2021b). Trem2 restrains the enhancement of tau accumulation and neurodegeneration by β-amyloid pathology. Neuron *109*, 1283–1301.e6. https://doi.org/10.1016/j.neuron.2021.02.010.

Lee, S.-H., Rezzonico, M.G., Friedman, B.A., Huntley, M.H., Meilandt, W.J., Pandey, S., Chen, Y.-J.J., Easton, A., Modrusan, Z., Hansen, D.V., et al. (2021a). TREM2-independent oligodendrocyte, astrocyte, and T cell responses to tau and amyloid pathology in mouse models of Alzheimer disease. Cell Rep. *37*, 110158. https://doi.org/10.1016/j.celrep.2021.110158.

Lee, Y., Morrison, B.M., Li, Y., Lengacher, S., Farah, M.H., Hoffman, P.N., Liu, Y., Tsingalia, A., Jin, L., Zhang, P.-W., et al. (2012c). Oligodendroglia metabolically support axons and contribute to neurodegeneration. Nature *487*, 443–448. https://doi.org/10.1038/nature11314.

Leegwater, P.A.J., Vermeulen, G., Könst, A.A., Naidu, S., Mulders, J., Visser, A., Kersbergen, P., Mobach, D., Fonds, D., van Berkel, C.G., et al. (2001). Subunits of the translation initiation factor eIF2B are mutant in leukoencephalopathy with vanishing white matter. Nat. Genet. *29*, 383–388. https://doi.org/10.1038/ng764.

Leng, K., Li, E., Eser, R., Piergies, A., Sit, R., Tan, M., Neff, N., Li, S.H., Rodriguez, R.D., Suemoto, C.K., et al. (2021). Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. Nat. Neurosci. *24*, 276–287. https://doi.org/10.1038/s41593-020-00764-7.

Li, W., Wang, X., Van Der Knaap, M.S., and Proud, C.G. (2004). Mutations linked to leukoencephalopathy with vanishing white matter impair the function of the eukaryotic initiation factor 2B complex in diverse ways. Mol. Cell Biol. *24*, 3295–3306. https://doi.org/10.1128/mcb.24.8.3295-3306.2004.

Lin, Y., Pang, X., Huang, G., Jamison, S., Fang, J., Harding, H.P., Ron, D., and Lin, W. (2014). Impaired eukaryotic translation initiation factor 2B activity spe-cifically in oligodendrocytes reproduces the pathology of vanishing white matter disease in mice. J. Neurosci. *34*, 12182–12191. https://doi.org/10.1523/jneurosci.1373-14.2014.

Louis, J., Magal, E., Takayama, S., and Varon, S. (1993). CNTF protection of oligodendrocytes against natural and tumor necrosis factor-induced death. Science *259*, 689–692. https://doi.org/10.1126/science.8430320.

Ma, X.-R., Zhu, X., Xiao, Y., Gu, H.-M., Zheng, S.-S., Li, L., Wang, F., Dong, Z.-J., Wang, D.-X., Wu, Y., et al. (2022). Restoring nuclear entry of Sirtuin 2 in oligodendrocyte progenitor cells promotes remyelination during ageing. Nat. Commun. *13*, 1225. https://doi.org/10.1038/s41467-022-28844-1.

Maluenda, J., Manso, C., Quevarec, L., Vivanti, A., Marguet, F., Gonzales, M., Guimiot, F., Petit, F., Toutain, A., Whalen, S., et al. (2016). Mutations in GLDN, encoding gliomedin, a critical component of the nodes of ranvier, are responsible for lethal arthrogryposis. Am. J. Hum. Genet. *99*, 928–933. https://doi.org/10.1016/j.ajhg.2016.07.021.

Marques, S., van Bruggen, D., Vanichkina, D.P., Floriddia, E.M., Munguba, H., Väremo, L., Giacomello, S., Falcão, A.M., et al. (2018). Transcriptional convergence of oligodendrocyte lineage progenitors during development. Dev. Cell *46*, 504–517.e7. https://doi.org/10.1016/j.devcel.2018.07.005.

Marques, S., Zeisel, A., Codeluppi, S., van Bruggen, D., Romanov, R.A., van Falcão, A.M., Mendanha Falcão, A., Xiao, L., Li, H., et al. (2016). Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. Science *352*, 1326–1329. https://doi.org/10.1126/science.aaf6463.

Mathys, H., Davila-Velderrain, J., Peng, Z., Gao, F., Mohammadi, S., Young, J.Z., Menon, M., He, L., Abdurrob, F., Jiang, X., et al. (2019). Single-cell transcriptomic analysis of Alzheimer's disease. Nature *570*, 332–337. https://doi.org/10.1038/s41586-019-1195-2.

Matsushima, G.K., and Morell, P. (2001). The neurotoxicant, cuprizone, as a model to study demyelination and remyelination in the central nervous system. Brain Pathol. *11*, 107–116. https://doi.org/10.1111/j.1750-3639.2001.tb00385.x.

Miron, V.E., Boyd, A., Zhao, J.-W., Yuen, T.J., Ruckh, J.M., Shadrach, J.L., van Wijngaarden, P., Wagers, A.J., Williams, A., Franklin, R.J.M., and Ffrench-Constant, C. (2013). M2 microglia and macrophages drive oligodendrocyte differentiation during CNS remyelination. Nat. Neurosci. *16*, 1211–1218. https://doi.org/10.1038/nn.3469.

Ozmen, L., Albientz, A., Czech, C., and Jacobsen, H. (2008). Expression of transgenic APP mRNA is the key determinant for beta-amyloid deposition in PS2APP transgenic mice. Neurodegener. Dis. *6*, 29–36. https://doi.org/10.1159/000170884.

Richards, J.G., Higgins, G.A., Ouagazzal, A.-M., Ozmen, L., Kew, J.N.C., Bohrmann, B., Malherbe, P., Brockhaus, M., Loetscher, H., Czech, C., et al. (2003). PS2APP transgenic mice, coexpressing hPS2mut and hAPPswe, show age-related cognitive deficits associated with discrete brain amyloid deposition and inflammation. J. Neurosci. *23*, 8989–9003. https://doi.org/10.1523/jneurosci.23-26-08989.2003.

Saher, G., Brügger, B., Lappe-Siefke, C., Möbius, W., Tozawa, R.i., Wehr, M.C., Wieland, F., Ishibashi, S., and Nave, K.-A. (2005). High cholesterol level is essential for myelin membrane growth. Nat. Neurosci. *8*, 468–475. https://doi.org/10.1038/nn1426.

Sanz-Rodriguez, M., Gruart, A., Escudero-Ramirez, J., de Castro, F., Delgado-García, J.M., Wandosell, F., and Cubelos, B. (2018). R-Ras1 and R-ras2 are essential for oligodendrocyte differentiation and survival for correct myelination in the central nervous system. J. Neurosci. *38*, 5096–5110. https://doi.org/10.1523/jneurosci.3364-17.2018.

Schirmer, L., Velmeshev, D., Holmqvist, S., Kaufmann, M., Werneburg, S., Jung, D., Vistnes, S., Stockley, J.H., Young, A., Steindel, M., et al. (2019). Neuronal vulnerability and multilineage diversity in multiple sclerosis. Nature *573*, 75–82. https://doi.org/10.1038/s41586-019-1404-z.

Shen, K., Reichelt, M., Kyauk, R.V., Ngu, H., Shen, Y.-A.A., Foreman, O., Modrusan, Z., Friedman, B.A., Sheng, M., and Yuen, T.J. (2021). Multiple sclerosis risk gene Mertk is required for microglial activation and subsequent remyelination. Cell Rep. *34*, 108835. https://doi.org/10.1016/j.celrep.2021.108835.

Simons, M., and Nave, K.-A. (2016). Oligodendrocytes: myelination and axonal support. Cold Spring Harb. Perspect. Biol. 8, a020479. https://doi.org/10.1101/cshperspect.a020479.

Srinivasan, K., Friedman, B.A., Etxeberria, A., Huntley, M.A., van der Brug, M.P., Foreman, O., Paw, J.S., Modrusan, Z., Beach, T.G., Serrano, G.E., and Hansen, D.V. (2020). Alzheimer's patient microglia exhibit enhanced aging and unique transcriptional activation. Cell Rep. 31, 107843. https://doi.org/10.1016/j.celrep.2020.107843.

Stankoff, B., Aigrot, M.-S., Noël, F., Wattilliaux, A., Zalc, B., and Lubetzki, C. (2002). Ciliary neurotrophic factor (CNTF) enhances myelin formation: a novel role for CNTF and CNTF-related molecules. J. Neurosci. 22, 9221–9227. https://doi.org/10.1523/jneurosci.22-21-09221.2002.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive integration of single-cell data. Cell 177, 1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

Trapp, B.D., and Nave, K.-A. (2008). Multiple sclerosis: an immune or neurodegenerative disorder? Annu. Rev. Neurosci. 31, 247–269. https://doi.org/10.1146/annurev.neuro.30.051606.094313.

Tyler, W.A., Gangoli, N., Gokina, P., Kim, H.A., Covey, M., Levison, S.W., and Wood, T.L. (2009). Activation of the mammalian target of rapamycin (mTOR) is essential for oligodendrocyte differentiation. J. Neurosci. 29, 6367–6378. https://doi.org/10.1523/jneurosci.0234-09.2009.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26, 873–881. https://doi.org/10.1093/bioinformatics/btq057.

Wu, Y.E., Pan, L., Zuo, Y., Li, X., and Hong, W. (2017). Detecting activated cell populations using single-cell RNA-seq. Neuron 96, 313–329.e6. https://doi.org/10.1016/j.neuron.2017.09.026.

Yoshiyama, Y., Higuchi, M., Zhang, B., Huang, S.-M., Iwata, N., Saido, T.C., Maeda, J., Suhara, T., Trojanowski, J.Q., and Lee, V.M.-Y. (2007). Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model. Neuron 54, 343–344. https://doi.org/10.1016/j.neuron.2007.03.022.

Yuen, T.J., Silbereis, J.C., Griveau, A., Chang, S.M., Daneman, R., Fancy, S.P.J., Zahed, H., Maltepe, E., and Rowitch, D.H. (2014). Oligodendrocyte-Encoded HIF Function Couples Postnatal Myelination and White Matter Angiogenesis. Cell 158, 383–396.

Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keeffe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA-Sequencing Transcriptome and Splicing Database of Glia, Neurons, and Vascular Cells of the Cerebral Cortex. J. Neurosci. 34, 11929–11947.

Zhang, Y., Sloan, S.A., Clarke, L.E., Caneda, C., Plaza, C.A., Blumenthal, P.D., Vogel, H., Steinberg, G.K., Edwards, M.S.B., Li, G., et al. (2016). Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. Neuron 89, 37–53. https://doi.org/10.1016/j.neuron.2015.11.013.

Zhou, Y., Song, W.M., Andhey, P.S., Swain, A., Levy, T., Miller, K.R., Poliani, P.L., Cominelli, M., Grover, S., Gilfillan, S., et al. (2020). Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. Nat. Med. 26, 131–142. https://doi.org/10.1038/s41591-019-0695-9.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
Chromium™ Single Cell 3′ Library & Gel Bead Kit v2	10X Genomics	PN-120237
Chromium Single Cell A Chip Kit	10X Genomics	PN-120236
Neural Tissue Dissociation Kit (P)	Miltenyi Biotech	130-092-628
Bioanalyzer High Sensitivity DNA kit	Agilent Technologies	5067–4626
Qubit dsDNA HS Assay Kit	ThermoFisher Scientific	Q33230
Deposited data		
TauP301S scRNA-Seq dataset	This paper	GSE180041
Lysolecithin scRNA-Seq dataset	This paper	GSE182846
OligoLandscape Web Resource	This paper	http://research-pub. gene.com/OligoLandscape/
Experimental models: Organisms/strains		
Prnp-MAPT*P301S	(Yoshiyama et al., 2007)	N/A
hMAPT-P301L mice (pR5-183)	(Götz et al., 2001)	M/A
PS2APP mice	(Richards et al., 2003)	N/A
TauPS2APP mice	(Grueninger et al., 2010; Lee et al., 2021a, 2021b)	N/A
TauPS2APP; Trem2KO mice	(Lee et al., 2021a, 2021b)	N/A
Mouse: C57BL-6	Charles River Hollister	N/A
Oligonucleotides		
smFISH oligonucleotides, see Table S4	This paper	N/A
Software and algorithms		
Ingenuity Pathway Analysis	Qiagen	Version 01-18-06
Zen	Carl Zeiss	Zen 2.3 SP1
Seurat	Stuart et al. (2019)	Version 3.2.0
Core Dataset Integration	This paper	10.5281/zenodo.6784453
Other		
Actinomycin D	Millipore-Sigma	A1410
Calcein Violet AM	ThermoFisher Scientific	C34858
Propidium iodide	ThermoFisher Scientific	P1304MP
l-a-lysophosphatidylcholine	Sigma-Aldrich	L4129
Cuprizone	TCI America	B0476

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact Shristi Pandey (pandey.shristi@gene.com).

Materials availability
This study did not generate new unique reagents.

Data and code availability

- Datasets have been deposited in the Gene Expression Omnibus (GEO) repository, accession numbers GEO: GSE180041 and GEO: GSE182846.
- Original code related to core data integration has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

All protocols involving animals were approved by Genentech's Institutional Animal Care and Use Committee, in accordance with guidelines that adhere to and exceed state and national ethical regulations for animal care and use in research. All mice were maintained in a pathogen-free animal facility under standard animal room conditions (temperature 21 ± 1°C; humidity 55%–60%; 12h light/dark cycle). Mice of both sexes were used. The following mouse strains were used in this study:

- C57BL6 (Charles River Hollister)
- B6; C3-Tg (Prnp-MAPT*P301S) PS19Vle/J.

For GSE182846, 3–4 months old mice were treated with 1% lysolecithin (l-a-lysophosphatidylcholine, Sigma-Aldrich) as follows. Mice were anesthetized with isoflourane, secured to a stereotactic frame and treated with Meloxicam prior to surgery. Using a drill with a burr tip size of <0.5 mm in diameter, a small hole was drilled at the stereotaxic coordinates 1.2 mm posterior and 0.5mm lateral relative to Bregma for unilateral injections into the right striatum. A 10μL Hamilton syringe with a pulled glass pipet containing 1% lysolecithin (l-a-lysophosphatidylcholine, Sigma-Aldrich) solution was inserted into the drilled hole, 1.63 mm deep into the cortex (measured from the top of the skull). The injection was slowly infused at a speed of 0.35μL/min. The incision was closed and animals allowed to recover in a heated blanket before being returned to their cage.

For GSE180041, 9-months old mice were perfused in ice-cold PBS for preparation of single cell suspensions.

For multiplexed *in situ* hybridization experiments, 3–4 months old mice were perfused with ice-cold PBS, followed by dissection of their brains.

METHOD DETAILS

Preparation of single-cell suspensions from the lysolecithin treated animals

Mice were perfused with cold PBS and the brain was submerged into cold cutting solution (high magnesium choline replacement ACSF) pumped with oxygen on a vibratome. Using the vibratome, 150μm thick coronal sections were cut and corpus callosum lysolecithin induced lesion tissue were micro-dissected out under a dissection scope. The tissue was then dissociated with Miltenyi Neural Dissociation Kit (P) according to the manufacturer's instructions. After being enzymatically dissociated, the cells are passed through a cell strainer to clean up debris. Cells were then resuspended in a 23% Percoll gradient and spun at 1000 g for 20 min to clean up myelin debris. For single-cell RNA-seq, the cell pellets were then resuspended in Hibernate-A (ThermoFisher) containing propidium iodide and Calcein AM for the selection of live cells using flow cytometry.

Preparation of single-cell suspensions from the TauP301S animals

9-month-old non-transgenic (n = 3) or TauP301S^{het} (n = 6) mice were perfused with cold PBS and the hippocampi were immediately sub-dissected. Single cell suspensions were prepared from the hippocampi as described in (Lee et al., 2021a). Briefly, hippocampi were chopped into small pieces and dissociated with enzyme mixes in Neural Tissue Dissociation Kit (P) (Miltenyi 130-092-628) in the presence of actinomycin D. After dissociation, cells were resuspended in Hibernate A Low Fluorescence medium (Brainbits) containing 5% FBS, with Calcein Violet AM (Thermo Fisher C34858) and propidium iodide (Thermo Fisher P1304MP). Flow cytometry was used to sort and collect live single-cell suspensions for the single-cell RNA-seq study.

Single-cell RNA-seq library preparation and sequencing

Sample processing and library preparation was carried out using the Chromium Single Cell 3′ Library and Gel Bead Kit v2 (10X Genomics) according to manufacturer's instructions. 6000–10000 cells were used for GEM generation, and libraries were sequenced with HiSeq 4000 (Illumina).

In situ hybridization, imaging and image analysis

For Figure 4, spatially resolved, multiplexed *in situ* RNA detection and analysis was performed using the automated Rebus Esper spatial omics platform (Rebus Biosystems, Inc., Santa Clara, CA). By intersecting a list of OL lineage, DAO, microglial, astrocytic, endothelial and neuronal candidate cell type markers generated by scRNA-seq, suitability for probe design, including gene length

and relative abundance, and design constraints for compatibility with the Rebus Esper spatial omics platform using proprietary software, we generated the following gene probe panel: OL Lineage markers (Olig2, Mog, Myrf), OPC markers (*Pdgfra*, *Cspg4*), COP (*Bmp4*), MOL2 markers (*Klk6*, *Hopx*), MOL5/6 markers (*Il33*, *Jph4*), DAO markers (*C4b*, *Serpina3n*, *Parp3*, *Steap3*, *Cdkn1a*, *Fosb*, *Egr1*, *Klf4*, *Snhg1*, *Irf7*, *Bst2*), microglia/DAM (*Tmem119*, *Spp1*, *Gpnmb*), astrocytes (*Aqp4*), Neurons (*Snap25*, *Syt1*) and endothelial cells (*Pecam1*) (Table S4). Experiments using the Rebus Esper spatial platform were performed as follows. Mice were perfused with ice-cold PBS, followed by dissection of their brains. Brains were embedded in OCT (Sakura Tissue-Tek O.C.T.), snap-frozen on dry ice and stored at −80°C. Fresh frozen brain tissue sections (5 μm) were cut on a cryostat, mounted on the functionalized coverslips and fixed for 10 min with 4% paraformaldehyde (Alfa Aesar, Cat#43368) in PBS at room temperature, rinsed twice with PBS at room temperature and stored in 70% ethanol at 4°C before use. The sample section on the coverslip was assembled into a flow cell, which was then loaded onto the instrument. The hybridization cycles and imaging were done automatically under the instrumental control software.

Briefly, primary probes for all target genes were initially hybridized for 6 h and probes not specifically bound were washed away. Readout probes labeled with Atto532, Atto594 and Atto647N dyes were then hybridized, washed, counterstained with DAPI and then imaged with an Andor sCMOS camera (Zyla 4.2 Plus, Oxford Instruments) through 20xC, 0.45NA dry lens (CFI S Plan Fluor ELWD, Nikon) with 365nm LED for DAPI and 532nm, 595nm and 647nm lasers configured for SAO imaging. Single Z-planes with 2.8μm depth of field were acquired for each field of view. Using the Rebus Esper image processing software, the raw images were reconstructed to generate high-resolution images (equivalent or better than images obtained with a 100× oil immersion lens). RNA spots were automatically detected to generate high fidelity RNA spot tables containing *xy* positions and signal intensities. Nuclei segmentation software based on StarDist (Fazeli et al., 2020) (identified individual cells by finding nuclear boundaries from DAPI images. The detected RNA spots were then assigned to each cell using maximum distance thresholds. The resulting cell × feature matrix contains gene counts per cell along with annotations for cell location and nuclear size.

QUANTIFICATION AND STATISTICAL ANALYSIS

Raw data analysis

When possible, scRNA-seq FASTQ files were downloaded from public repositories (Gene Expression Omnibus, Short Read Archive, Synapse). In most cases, raw datasets were analyzed with an in-house pipeline. Briefly, reads were demultiplexed based on perfect matches to the expected cell barcodes. Transcript reads were aligned to the mouse reference genome (GRCm38) or the human reference genome (GRCm38) using GSNAP (Wu and Nacu, 2010). Only uniquely mapping reads were considered for downstream analysis. Transcript counts for a given gene were based on the number of unique UMIs (up to one mismatch) for reads overlapping exons in sense orientation. Cell barcodes from empty droplets were filtered by requiring a minimum number of detected transcripts. Sample quality was further assessed based on the distribution of per-cell statistics, such as total number of reads, percentage of reads mapping uniquely to the reference genome, percentage of mapped reads overlapping exons, number of detected transcripts (UMIs), number of detected genes, and percentage of mitochondrial transcripts.

In the case of GSE157827, FASTQ files were downloaded and converted into matrices of expression counts using the cellranger software provided by 10X Genomics. Finally, for GSE129788, GSE147528 and syn17115627, the author provided gene expression matrices were directly used for the analysis. Mouse and human datasets included in the study are listed in Table S1.

Mouse scRNA-seq meta-analysis

A total of eight mouse datasets were analyzed spanning 5 AD models and 3 MS models representing datasets collected from public repositories and in-house datasets (Table S1). Five datasets were included in the core integration analysis. Integrated atlases were generated and analyzed using the following steps: 1) processing and filtering of each individual scRNA-seq dataset, 2) integration along datasets to generate a single atlas 3) Transfer of labels to the remaining datasets.

Processing and filtering individual scRNA-seq datasets

For public datasets, we analyzed the raw data using the raw data processing described above. For each dataset, we filtered for low quality cells with <200 measured genes and a high percentage of mitochondrial DNA contamination (<5%). After filtering, each dataset was normalized to log (CPM/10⁵), the 2000 most variable genes in each dataset identified and the expression levels of these genes were scaled before performing PCA and clustering in the variable gene space. These steps were performed using functions implemented in the Seurat package (NormalizeData, FindVariableFeatures, ScaleData, RunPCA, FindNeighbors, FindClusters, RunUMAP) with some modification to default parameters as needed. OL lineage cells from each dataset were identified based on the expression of a set of markers of the OL lineage obtained from the Barres lab (Zhang et al., 2014). Cells with high expression of these genes were then subset out of the Seurat Object to create a new OL-lineage enriched Seurat object for each dataset.

Dataset integration for generation of mouse OL lineage atlas

Before dataset integration, we imported the aforementioned filtered OL-lineage enriched Seurat objects across multiple datasets. Next, each dataset now containing only the OL lineage cells were normalized (function NormalizeData, method = 'LogNormalize, scale.factor = 10,000) and scaled to regress out mitochondrial genes and batch effects. For each dataset, 2500 variable genes were computed using function (FindVariableFeatures, selection.method = 'vst'). Pairwise integration anchors were computed between

pairs of datasets. In the case of Cuprizone and PS2APP datasets, batch effects were observed within the dataset. Hence, for these datasets, each batch was treated as a separate dataset during the integration process. The resulting anchorset was then used to integrate all the cells into a single atlas. The integrated assay was then used to compute PCA, UMAP and clustering outputs. Next, we identified the distinct cell types based on pre-existing marker genes for OL lineage subtypes as well as an independent analysis of pseudobulk profiles obtained from the clustering output (see section on pseudo-bulk analysis of sc/snRNA-seq datasets). Upon clustering the dataset, differential abundance statistics were calculated using EdgeR.

Pseudo-bulk analysis of sc/snRNA-seq datasets

Pseudo-bulk expression profiles were derived from single cell datasets by aggregating cells of a given sample of the same cell type separately such that a total of sc possible pseudobulks were generated for s samples and c cell types. If a particular pseudobulk contained fewer than 10 cells, these profiles were discarded such that the actual number of pseudobulks was always less than the theoretical maximum number of pseudobulks, sc. For each pseudobulk profile, raw counts were generated by adding the total number of UMIs for each gene across all the cells belonging to a particular sample and cell type. This resulted in a gene-by-pseudobulk count matrix which was normalized using the estimatedSizeFactors function from DEseq2. These normalized pseudobulk profiles were then used for differential gene expression analysis between groups of interest and for calculation and visualization of gene set scores.

In formal notation, let n_{ij} be the raw UMI count of gene i in cell j and s_j and c_j be the sample and cell type, respectively of cell j. The pseudobulks for all cells are represented by pairs (s, c) where s is a sample and c is a cell type. Therefore, the pseudo-bulk count matrix B, with rows indexed by genes and columns indexed by (s, c) pairs is defined by:

$$B_{i, sc} = \sum_{j:(s_j, c_j) = (s,c)} n_{ij}$$

Matrix B is size-factor normalized and used for downstream differential gene expression with voom.

For each mouse dataset incorporated into the study, a methodology was applied to identify DE genes that were likely artifacts due to contamination from free RNA from other cell type compartments in the dataset as described in (Lee et al., 2021a). Briefly, to identify DEs contributed by other cell types, the following steps were applied:

1. Within the broader single cell RNA-seq dataset, a "two-compartment" pseudobulk matrix C was computed. The two compartments were Oligodendrocytes (denoted here by C_{OL}) and a second "other" compartment for all of the other cell types (denoted here by C_{Other}). In the notation above, the two-compartment pseudobulks are the pairs (s,c) where s is a sample and c is either the cell type of interest C_{OL} or C_{Other}, representing a bulk of all the other cells. Conceptually "other" represents the aggregate of all of the other cells, and is a model for the source of free RNA.

2. Two voom + limma DE analysis was performed on the matrix C, using linear models $\sim sampleID + compartment$ ("compartment model") and $\sim sampleID + compartment + compartment : genotype$ ("interaction model").

3. Prediction of free RNA contamination is made in two steps. In the first step, a gene is flagged as likely free RNA contamination if the *compartment* term from the "compartment model" has adjusted $P \leq 0.05$ and the "other" compartment is at least 6-fold higher than the C_{OL} compartment. The 6-fold cutoff can be tuned to other values, but was selected empirically for this study because the *compartment* effect sizes were bimodal for many comparisons of interest, and this cutoff cleanly separated the modes. Conceptually, this step involves flagging genes whose expression is (at least 6-fold) higher in the "other" cell types than in C_{OL}.

4. In the second step, the interaction term of the interaction model is examined. If this term is significant (adjusted $P \leq 0.05$) and the direction of this effect is in the same direction as the original pseudobulk differential expression, then the gene in unflagged. Conceptually, this step means that even though there is higher expression in the "other" compartment, which might contribute some signal to the C_{OL} compartment, the fold-change within the C_{OL} compartment is too large to explain solely by contamination from the "other" compartment, and therefore is likely explained by true DE within the C_{OL} compartment.

5. Genes flagged as "predictedFreeRNAContam" in greater than 2 datasets were removed from further analysis.

Differential expression (DE) analysis from pseudobulk sample

A number of different gene sets were generated in this study: 1) Gene markers associated with resting OL subtypes (Figure S1D); 2) Gene markers associated with disease-associated oligodendrocytes (Cluster agnostic) (Figure S3A) 3) Gene markers associated with each disease-associated oligodendrocyte state (Figure S4A).

For the generation of the gene markers associated with resting oligodendrocyte states, we performed differential gene expression analysis between each resting oligodendrocyte lineage cluster to the rest of the oligodendrocytes. For the generation of gene signatures associated with the disease-associated oligodendrocytes, we performed differential expression analysis between a pseudobulk of all disease-associated oligodendrocyte clusters and compared them to resting oligodendrocytes. Specifically, this included the following comparison pairs: MOL5/6_DA1 and MOL5/6, MOL5/6_DA2 and MOL5/6, MOL5/6_IFN and MOL5/6, MOL2_DA1 and MOL2, MOL2_IFN and MOL2 (Figure S2B). All differential expression analysis was performed with *voom* using dataset as the blocking variable.

Each comparison yielded a gene set that was either uniquely differential in a particular state or differentially expressed between multiple states, albeit at varying fold change values. For the shared genes, we categorized them as DA1, DA2 or IFN-enriched genes if their Log-fold change was highest in one out of the three states. If the fold change of a gene was comparable between multiple states within a 30% window of the maximum log fold change, the gene was categorized as being enriched in multiple states.

For pathway analysis, all differentially expressed genes between disease-associated states and resting states of mature oligodendrocytes were imported into Ingenuity Pathway Analysis (IPA) software and were subjected to IPA core expression analysis. Genes from the integrated analysis that met the cutoff values (Log2Fold change >0.6, FDR <0.05) were included in the analysis. Expr Log Ratio was used to determine the directionality (z-scores) in the analysis. Reference set of Ingenuity Knowledge base genes were used for p value calculations.

Gene set analysis
Gene Sets used in the Figures are as follows:

- Figures S1B and S1D: Full gene marker list in Table S1 Cell type marker genes for the resting oligodendrocyte subtypes (selected genes chosen for visualization). See Table S2 (Column name: OL.lineage.subtypes)
- Figure S3A: Full gene marker list with FDR and Log Fold Changes in Table S2 listed as Coarse DAO activation
- Figures S4A and S4B: Full gene marker list with FDR and Log Fold Changes in Table S2, listed as Fine DAO Activation. Genes also annotations in the columns 'DAO_fine_up' and 'DAO_fine_down' for whether each gene is classified as MOL5/6_DA1, MOL5/6_DA2 or MOL5/6_IFN. Top 40 selected genes are shown in the heatmap and the feature plot.
- Figures 6D, 6E, and S9: Genes were generated by performing DGE between 1) hOligo2 derived from control and MS patients 2) hOligo1 derived from control and MS patients across integrated MS datasets. Gene list and associated statistics in Table S3 under 'Integrated MS datasets' columns.
- Figures 7D and S10. Genes were generated by performing DGE between hOligo2 derived from high pathology and low pathology individuals across integrated AD datasets and selecting genes that are DE with LFC >= 1.5 and FDR <= 0.05. Gene list and associated statistics in Table S3 under 'integrated AD meta-analysis' columns

For a given gene set, gene set scores were calculated as follows. Gene expression values were log-transformed as Log2(normCount+1). The average gene set score for a sample was calculated as the average over all genes in the gene set. The average log-expression values of the controls were then subtracted out of the average of each sample to generate control-centered gene expression values. When including both upregulated and downregulated genes, a signed average was computed by weighting all the up genes by +1 and down genes by −1 to capture both changes in a single score.

Human scRNA-seq meta-analysis
A human disease-state oligodendrocyte atlas was constructed using snRNA-seq datasets generated with 10X technology for AD and MS. A total of nine (three MS and six AD) datasets were incorporated into the study and are listed in Table S1. Individual datasets were processed according to the description provided in Raw Data Processing except for Leng et al., Absinta, et al. and Cain et al., where processed count matrices provided by the authors were used for data analysis. Each dataset was analyzed separately to identify cells of the oligodendrocyte lineage as described before in the mouse meta-analysis. Human OL lineage cells were identified via the expression of canonical marker genes for OL lineage except for Leng et al., and Cain et al., where the author-provided annotations were used to subset out the clusters of oligodendrocyte lineage cells. Gene sets used to identify oligodendrocyte lineage cells were obtained from a sorted human RNA-seq dataset generated by the Barres lab (Zhang et al., 2016)

The extracted oligodendrocyte lineage cells from each dataset were analyzed separately to identify oligodendrocyte subtypes using aforementioned methodologies to filter, normalize, scale and cluster cells. Cells with <300 measured genes and >5% mitochondrial counts were removed from each dataset. The resulting filtered dataset containing just the oligodendrocyte lineage cells were normalized to log (CPM/10^5), the 2500 most variable genes were identified and the expression levels of these genes were scaled before PCA. Depending on the dataset, 10–15 most significant PCs were then used for graph clustering (resolution = 0.5) and UMAP dimensionality reduction. All steps were performed using Seurat package. Preliminary marker genes for each cluster were identified using FindAllMarkers or FindMarker function. Individual clusters were merged into larger clusters based on a dendrogram built on the variable gene set if no markers were able to distinguish a cluster from the rest of the dataset at an avg_LogFC >0.75.

To define resting human oligodendrocyte subtypes, cross-species integration analysis was performed with resting oligodendrocytes derived from control individuals across both species. Human oligodendrocytes were sub-clustered into hOligo1 and hOligo2 based on their co-clustering with MOL2 and MOL5/6. Human hOligo1 and hOligo2 markers were derived by performing a subtype-specific DE analysis using pseudobulk profiles obtained from the clustering output of this analysis (Figure 5) and used to annotate all subsequent human data.

For the MS datasets, we integrated the cells obtained from the three datasets into a single atlas by using an anchorset that was derived by identification of pairwise integration anchors on the genes that are commonly variable across the datasets. The integrated assay was used to compute PCA, UMAP and clustering outputs. Clusters were annotated using a gene set score of the markers derived from analysis of control oligodendrocytes. The clusters recovered from the individual analysis of MS datasets and the

integrated analysis were largely identical, except for the identification of hOligo4 which were only identified in the Schirmer *et al* and Absinta et al. in the individual analysis (Figure S8C). We also found a small cluster of cells that expressed markers of both microglia and oligodendrocytes that were ruled to be potential doublets and removed from further analysis. Another small cluster of cells expressing both oligodendrocyte and astrocyte markers were also identified in multiple datasets and removed as potential doublets. Upon clustering the dataset, differential abundance statistics were calculated using EdgeR. The clustering outputs for cells obtained from Jäkel et al. were also compared to the oligodendrocyte clustering annotations described by the authors in Jäkel et al. We found that hOligo1 largely mapped to Oligo1 and Oligo5 whereas hOligo2 largely mapped to Oligo2, Oligo3, Oligo4 and Oligo6.

For the AD datasets, we integrated the cells obtained from the six datasets into a single atlas by using an anchorset that was derived by identification of pairwise integration anchors on the genes that are commonly variable across both datasets. The integrated assay was used to compute PCA, UMAP and clustering outputs. Clusters were annotated using markers of human oligodendrocytes derived from the analysis of resting oligodendrocytes. The clusters recovered from the individual analysis of AD datasets and the integrated analysis were largely identical (Figure S10B).

Differential expression (DE) analysis from pseudobulk sample to obtain an MS and AD-associated oligodendrocyte activation signature

To identify AD and MS-associated transcriptional shift in oligodendrocyte subtypes, we aggregated cells of a given subtype and the same patient into pseudobulk profiles. For MS datasets, we created pseudobulk profiles from the cluster definitions obtained via the integrated analysis of MS datasets. Then, using dataset as a blocked variable and controlling for age, sex and post-mortem interval as covariates, we performed DE between MS and controls within each oligodendrocyte subtype compartment.

For AD datasets, we created pseudobulk profiles from oligodendrocyte subtypes obtained from the integrated analysis of AD datasets – stratifying patients into high pathology (high levels of amyloid plaques and high tau pathology) and low pathology (low to no amyloid plaques and low tau pathology) individuals. In most cases, low pathology individuals were defined as patients who were in Braak Stage 0, I and II and CERAD score 4 (No AD) and high pathology individuals were defined as patients in Braak Stage V and VI and CERAD score 1 (Definite AD) and 2 (Probable AD). Subtype-specific DE between high and low pathology individuals were performed using dataset as a blocking variable and controlling for age, sex, post-mortem interval as covariates. Because CERAD score was not available for all datasets, only those patients containing information on both ceradsc and Braak Stage were included in the DE analysis. During gene set visualization, in the datasets where CERAD score was lacking, patients were grouped into low and high pathology groups based solely on Braak Stage.

Analysis of ISH data

The expression matrix obtained from Rebus Biosystems was filtered to remove cells with less than 5 or greater than 120 spot counts, cells with nucleus perimeter less than 8 microns and greater than 50 microns. To remove doublets resulting from spot misassignment in the spatial transcriptomics dataset, cells containing more than 2 counts of transcripts belonging to orthogonal cell types were removed from further analysis. For example: neurons and microglia, astrocytes and microglia, microglia and oligodendrocytes, microglia and OPCs, etc. All downstream clustering analysis of the dataset was performed using Seurat setting all 31 genes as variable features in the analysis. The resulting clusters were annotated based on expression of the markers selected for spatial transcriptomics. Cells from the corpus callosum were computationally selected using the average expression of the oligodendrocyte gene module (*Olig2*, *Myrf*, *Mog*) and neuronal gene module (*Syt1* and *Snap25*). While *Syt1* expression was largely limited to regions outside of the corpus callosum, the oligodendrocyte gene module expression was largely limited to inside the corpus callosum. For each cell, the expression of the two gene sets were thresholded to assign each cell as 'Neuron', 'Oligodendrocytes' or 'Other'. For each cell, a local neighborhood (radius = 20μm) of cells were identified. If a majority of the cells within a neighborhood of a particular cell were assigned as 'Oligodendrocytes', the cell was labeled as being inside the corpus callosum. Otherwise, the cell was labeled as being outside of the corpus callosum. This roughly demarcated the corpus callosum cells from the rest of the surrounding tissue but led to some cells within the corpus callosum to be labeled as non-corpus callosum cells. These were manually annotated as corpus callosum cells using the points.in.polygon function in the library sp. This led to the identification of corpus callosum cells as shown (Figure S6C) which were used for downstream spatial analysis. To calculate the spatial proximity of cell populations to the lateral lesion site, a straight line going through the midline of the corpus callosum was used to scale the distance from the midline of the corpus callosum to the left or right end (Figure S6C). This scaled distance was then used to subdivide each corpus callosum into six equidistant bins. The proportion of cells in each bin was then calculated to determine the change in cellular proportions from varying distances from the lesion site (Figure 4F).