Single nuclei RNAseq stratifies multiple sclerosis patients into three distinct white matter glia responses

Will Macnair[1#], Daniela Calini[1#], Eneritz Agirre[2‖], Julien Bryois[1‖], Sarah Jäkel[3,4‖], Petra Kukanja[2], Nadine Stokar-Regenscheit[5], Virginie Ott[5], Lynette C. Foo[1], Ludovic Collin[1], Sven Schippling[1], Eduard Urich[1], Erik Nutma[6], Manuel Marzin[6], Sandra Amor[6], Roberta Magliozzi[7], Elyas Heidari[8], Mark D. Robinson[8], Charles ffrench-Constant[9*], Gonçalo Castelo-Branco[2,10*], Anna Williams[11*], Dheeraj Malhotra[1*]

[1] Neuroscience and Rare Diseases, Roche Innovation Center, Basel, Switzerland
[2] Laboratory of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 17177 Stockholm, Sweden
[3] Institute for Stroke and Dementia Research (ISD), Klinikum der Universität München, Ludwig-Maximilians Universität, München, Germany
[4] Munich Cluster for Systems Neurology (SyNergy), München, Germany
[5] Pathology Chapter, Roche Pharma Research and Early Development, Roche Innovation Center, Basel, Switzerland
[6] Department of Pathology, Amsterdam UMC-Location VUmc, Amsterdam, The Netherlands
[7] Department of Neuroscience, Biomedicine and Movement Science, University of Verona, Verona, Italy
[8] Department of Molecular Life Sciences and SIB Swiss Institute of Bioinformatics, University of Zurich, Winterthurerstrasse 190, 8057, Zurich, Switzerland
[9] Faculty of Medicine and Health Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ UK
[10] Ming Wai Lau Centre for Reparative Medicine, Stockholm node, Karolinska Institutet, 171 77 Stockholm, Sweden
[11] Centre for Regenerative Medicine, Institute for Regeneration and Repair, MS Society Edinburgh Centre for MS Research, The University of Edinburgh, Edinburgh BioQuarter, 5 Little France Drive, Edinburgh EH16 4UU, UK

* Equal contribution

‖ Equal contribution

Correspondence to Dheeraj Malhotra or Anna Williams or Gonçalo Castelo-Branco or Charles ffrench-Constant

Abstract

The lack of understanding as to the cellular and molecular basis of clinical and genetic heterogeneity in progressive multiple sclerosis (MS) has hindered the search for new effective therapies and biomarkers. Here, to address this gap, we analysed 740,000 single nuclei RNAseq profiles of 165 samples of white matter (WM) lesions, normal appearing WM, grey matter (GM) lesions and normal appearing GM from 55 MS patients and 28 controls. We find that gene expression changes in response to MS are highly cell-type specific in WM and GM lesions but are largely shared within an individual cell-type across lesions, following a continuum rather than discrete lesion-specific molecular programs. The major biological determinants of variability in gene expression in MS samples relate to individual patient effects, rather than to lesion types or other metadata. Using multi-omics factor analysis (MOFA+), we identify three subgroups of MS patients with distinct oligodendrocyte composition and WM glial gene expression signatures, suggestive of engagement of different pathological/regenerative processes. The discovery of these three patterns significantly advances our mechanistic understanding of progressive MS, provides a framework to use molecular biomarkers to stratify patients for best therapeutic approaches for progressive MS, and highlights the need for precision-medicine approaches to address heterogeneity among MS patients.

Introduction

Despite a number of approved, highly effective therapies for the relapsing-remitting phase of multiple sclerosis (MS), we lack such therapies specifically targeting mechanisms of neurodegeneration in progressive stages of the disease. Therapeutic strategies that have been tested in clinical trials include enhancing neuroprotection directly and enhancing remyelination, aiming to provide indirect neuroprotection by restoring metabolic support and saltatory conduction to the demyelinated axon[1]. However, in spite of promising preclinical data, such trials so far have not met their primary endpoint of improvement in clinical disability, even though subgroup analysis has shown some promise (e.g. MS-SMART[2], Opicinumab[3], Bexarotene[4], Clemastine[5]). This translational mismatch may result from the heterogeneity of the disease in people with MS. Within both primary and secondary progressive MS (PPMS and SPMS) clinical subtypes, there is a clear heterogeneity of clinical course, with some people with MS showing a slowly progressive disability phase, while others become very disabled very quickly. This diverse disease course is very difficult to predict at disease onset. Speculating that a heterogeneous neurodegenerative and/or neuroregenerative response to MS pathology between patients underlies these differing disease outcomes, we and others have, in previous work, identified cellular heterogeneity in MS using single nucleus transcriptomics, albeit in a limited number of patients and few pathological MS lesion types[6–8]. However, these studies had insufficient samples to characterise inter-patient heterogeneity of demyelinated lesions and intra-patient heterogeneity between different lesions. To address this critical gap, we performed a single nucleus RNA sequencing study on the most extensive cohort of MS patients to date (Fig. 1a), including both white matter (WM) and grey matter (GM) areas. Our goals were firstly to identify the basis of heterogeneity by comparing cell type-specific gene expression signatures and cellular compositions across WM and GM MS lesion types. Secondly, we sought to identify pathologically relevant ways of stratifying patients on the basis of this response, so as to better find and test potential therapies for progressive MS.

Cell type-specific transcriptional landscape in normal and MS brains

We profiled 173 WM and GM samples, resulting in pre-QC totals of >950,000 nuclei from 55 MS cases (mainly PPMS and SPMS) and 30 controls. After randomization of samples during library preparation and sequencing to minimise batch effects (including 1 sample of each lesion type and 1 control in each batch), then doublet removal, cell and sample QC (Methods), we obtained 740,023 single-nucleus transcriptomes from 165 QC passed samples, including

585,998 nuclei from 55 MS patients and 154,025 nuclei from 28 controls, profiled at a median depth of 4,194 nuclei/sample, 3,154 reads/nucleus and 1,702 genes per nucleus (Extended Data Figs. 1c-e, Supplementary Files 1 and 2). Our cohort was balanced for age, gender and post-mortem interval, with cases and controls having similar distributions (Supplementary File 1, Extended Data Fig. 1a,b). These included 62 WM lesions (21 active, 17 chronic active, 13 chronic inactive and 11 remyelinated; respectively AL, CAL, CIL, RL; see Methods for details of dissection), 17 adjacent normal-appearing white matter (NAWM) regions from MS patients, and 15 cortical hemisphere WM regions from non-neurological controls[9]. In addition, we profiled 39 subpial cortical GM demyelinated lesions (GML), 16 adjacent normal appearing grey matter (NAGM) regions from MS patients and 16 cortical GM tissues from controls, all defined as per classical neuropathology[10], thereby creating a comprehensive atlas of single-nuclei MS transcriptomes (Fig. 1a). We used Conos[11] for integration and clustering of the QC filtered nuclei followed by merging of highly similar clusters using SCCAF[12]. This identified 50 distinct batch-corrected cell type clusters (Fig. 1b,c, Methods) capturing all major cell types of the human cortical GM and WM, with heterogeneity within these, including 18 subtypes of cortical excitatory neurons (across layers 2-6), 7 of inhibitory neurons, 9 of oligodendrocytes, 3 of committed oligodendrocyte precursor (COP), 5 of astrocytes, 2 each of microglia and endothelial cells, and 1 each of oligodendrocyte precursor cells (OPC), pericytes, B and T cells (Extended Data Fig. 2a,b). Based on the expression of previously-described genes characterising oligodendroglia[13], and confirming the heterogeneity we have previously identified[6] (Fig. 1c,d), we allocated the committed oligodendrocyte precursors (COP) into 3 subgroups (A1, A2 and B) and the oligodendrocytes into 9 subgroups (A1-2, B1-4, C1-2 and D), with the A to B to C nomenclature reflecting a blueprint usually found during developmental maturation and the Oligo D subgroup characterised by expression of genes associated with cellular stress (Fig. 1c). We annotated distinctive states of oligodendrocyte subpopulations using gene module analysis (Fig. 1e). PAGA analysis[14] revealed a putative trajectory from OPC to Oligo A, via COP A and then COP B (Fig. 1f), and Oligo C1 and D occupy a position consistent with their being end states. Whether the interconnected Oligo B populations represent end states or intermediates in the pathway to Oligo C is unclear, as it is difficult to unambiguously determine direction from single nuclei data[15]. Conos-defined clusters showed high concordance with clusters identified using an orthogonal data integration method, Seurat v4.0.5+Harmony[16,17] (Extended Data Figs. 2c), and we found no cluster composed of nuclei captured only from individual patients, samples, lesion types or technical covariates, indicating that data integration was successful (Fig. 1b and Extended Data Figs. 2d).

Distinct cell type specific transcriptional responses in GM and WM lesions

Annotating major cell type clusters and subclusters using established cell lineage and *de novo* marker genes, we compared cell type-specific gene expression changes between WM lesions and control WM tissue, and between GM lesions and control GM tissue using a mixed model (glmmTMB[18]) fit to pseudobulk data, including age, sex and post-mortem interval (PMI) as possible confounding variables, and donor ID as a random component (Methods, Supplementary Note). The distribution of random model fits in each broad cell type showed a strong donor effect for many genes (Extended Data Figs. 3a,b). Nevertheless, we identified 2,914 DEGs in GM, and 4,440 differentially expressed genes (DEGs) in WM in total across all 9 CNS major cell types (Fig. 1g,h, Supplementary Files 3 and 4, and https://malhotralab.shinyapps.io/MS_broad/).

In GM lesions, all major CNS cell types showed significant transcriptional changes in NAGM and GML, with strongest effects in excitatory and inhibitory neurons (Fig.1g (top)). Notably, both excitatory and inhibitory neurons show many more DEGs (over 4 times more up and downregulated genes in inhibitory neurons) in GML compared to NAGM. These changes were most prominent in specific subpopulations of excitatory neurons, Ex_RORB_A and Ex_RORB_CUX2A, and parvalbumin-positive neurons, Inh_Pvalb_A (Extended Data Fig. 4a) consistent with prior observations of selective vulnerability of excitatory and inhibitory neuronal subpopulations in MS GM lesions[7,19].

Transcriptomic and pathway changes in excitatory projection neurons in MS were strongly linked to an upregulation of genes related to glutamate signalling (*GRIA1*, *GRIA2*, *GRIA4*, *GRIN2B*, *GRM1*, *GRM5*), glucose or cation homeostasis (*SLC2A12*, *SLC22A10*) with concurrent down regulation of specific ion channels (*SCN1A, SCN1B, SCN2B, SCN4B, KCNA1, KCNA2, KCNC1*) and oxidative phosphorylation (OXPHOS) genes (A*TP1A1, ATP1B1, NDUFB10, NDUFS3, UQCRH*) (Fig. 1h(top),i). Cell stress-related genes and pathways such as oxidative stress and heat-shock response genes were not affected, suggesting that glutamate excitotoxicity via both increased excitatory and decreased inhibitory tone might be a critical determinant of selective neuronal GM pathology, and therefore is a promising therapeutic intervention in progressive MS.

In WM lesions, astrocytes, oligodendroglia and microglia show the strongest changes in gene expression (Fig. 1g (bottom)). The majority of DEGs were perturbed only in one broad cell type indicating strong cell type specificity of perturbations across lesions. Oligodendrocytes showed considerably more down-regulated genes across lesions while DEGs in astrocytes,

microglia and OPCs showed equivalent changes in up- and down-regulated genes. The numbers of significant DEGs for neuronal, endothelial and pericyte populations were substantially smaller, likely reflecting reduced power due to lower abundance of these cell types in WM.

Pathway analyses in WM lesions confirmed our prior observations that expression of genes involved in interferon alpha and gamma responses varied across lesions, and showed opposite patterns in OPCs as compared to oligodendrocytes (Fig. 1h (bottom),j). This suggests that the lesion environment at these stages might lead OPCs, but not oligodendrocytes, to transition to a state incompatible with immune competence[20]. Genes involved in inflammation-related pathways were enhanced in astrocytes, microglia, oligodendrocytes and endothelial cells in WM lesions (Fig. 1h (bottom)).

To investigate whether genes implicated in MS genetic risk are enriched in cell type-specific disease gene programs in MS, we quantified the over-representation of MS GWAS genes in cell type-specific DEGs in WM and GM lesions using MAGMA[21]. As expected from previous work[22], we found significant enrichment of MS risk genes among genes up and downregulated in immune cells, which was present across all MS tissue, and in MS lesions in microglia (Fig. 1k). We also found significant enrichment of MS risk genes among genes upregulated in pericytes (across all MS lesion types), genes upregulated in endothelial cells (in AL, RL and GML), and genes upregulated in OPC/COPs in RL (Fig.1k, Supplementary File 5). Furthermore, in our earlier report[23], integrating cell type-specific eQTLs with MS GWAS, we found cell type-specific roles for several MS risk genes; some of which show significantly dysregulated expression in the implicated cell-types in this snRNA-seq dataset, for example in excitatory neurons (*STAT4, SLC12A5, ANKRD55, ARHGAP27*), in inhibitory neurons (*SLC12A5, MTFR1L, H2AFX*), in oligodendrocytes (*ASPHD1, DBF4B, AHI1, RPS6KA4*), and in astrocytes (*TGFBR3, NR1H2, MXD3*) in GM and WM lesions, providing support for the role of these cells in MS disease risk and progression.

For many cell type-specific DEGs in the WM lesions, we observed ‘u’/’n’-shaped profiles of transcriptional changes along the pathological category of the lesion, i.e. that NAWM lesions showed small fold changes relative to control WM, increasing to the largest fold changes in AL and CAL, then decreasing in size in CIL and RL (pattern 1 in OPCs and microglia (’n’), patterns 6 in astrocytes and 5 in microglia (‘u’) Fig. 2a, Methods). Furthermore, we find that within each glial broad type and their subclusters, the majority of DEGs are shared across lesions (Fig 1g (bottom), Extended Data Fig. 4b). Astrocyte A, Microglia A, Oligo B4 and Oligo C1 subpopulations showed strongest transcriptional changes across lesions with a similar ‘n’-shaped pattern of increase in gene expression from NAWM to AL to CAL and then decrease

in CIL and RL (Extended Data Fig. 4b). These comparisons therefore show distinct changes associated with MS, while at the same time showing that neuropathologically-defined lesion categories represent mostly a continuum of transcriptional pathology, independent of cell type, rather than distinct gene expression programs.

Biological determinants of transcriptional heterogeneity in MS brains

We found in our analysis of transcriptional responses above that the cell-type specific changes are largely shared across lesions, and that for many genes, patient identity is a much stronger driver of variability in cell type-specific gene expression than lesion type in both WM and GM (Extended Data Fig. 3a,b). This led us to explore transcriptional heterogeneity in lesions within individual patients. Leveraging the unique strength of our study, namely many patients from whom we have samples with multiple different lesion types, we examined the expression patterns of genes that either show significant disease effect, or are highly variable between patients (Extended Data Fig. 3a,b). In both WM samples (Fig. 2b,c) and GM samples (Fig. 2d,e), we find strong evidence of coherent cross-cell type (tissue level) transcriptional heterogeneity between patients. Within an individual patient, the gene expression profiles are remarkably similar across multiple lesions and NAWM, while different patients show distinct yet shared transcriptional profiles across their lesions (Fig. 2b-e). We conclude that the donor ID (patient) is a much stronger driver of variability in cell type-specific gene expression than lesion type in both WM and GM.

To characterise this patient-driven transcriptional heterogeneity in our dataset we used MOFA+, a computational method developed for identifying low-dimensional representations of variation across multiple data modalities measured in the same samples[24]. In this study, we took the different cell types to be the different modalities. This allows us to identify responses that are coherent across samples, across multiple cell types simultaneously, and which may have cell type-specific responses (Extended Data Fig. 5a). MOFA+ does this by finding factors that seek to explain the variability in the input data (intuitively similar to PCA) across samples, and which correspond to coordinated tissue-level responses, even though the genes identified for each cell type are distinct. For example, a factor corresponding to remyelination might involve myelination genes in oligodendrocytes, debris-clearance genes in microglia, and metabolic support genes in astrocytes. These factors can then suggest possible ways to stratify patients.

It is recommended to first restrict to relevant genes before using MOFA+. For each cell type, we therefore selected genes with evidence of an MS effect and/or a donor effect (Extended Data Fig. 3a,b), to capture both consistent MS pathology and patient-patient variability; this results in sets of genes that are largely distinct for each cell type, with a small number of common genes. MOFA+ identified 5 factors each in both GM and WM samples that explained at least 5% of variability for some celltype (Extended Data Fig. 5b,c), with each factor describing one axis of variation in MS. If the factor explains variance in multiple cell types simultaneously, then the factor describes a response to MS that is coherent across multiple cell types, although the programs may be distinct for each celltype.

In GM samples, for all factors except factor 1, there was considerable overlap between the distributions of control and MS samples (Extended Data Fig. 5d). Notably, MS diagnosis and lesion pathology explained 64% variability in factor 1 (Extended Data Fig. 5b) indicating that factor 1 gene expression could robustly distinguish GM pathology (NAGM and GML) from control GM. Factor 1 explains equivalent variability in expression of genes across all major CNS cell types (Extended Data Fig. 5b) and high factor 1 (and GM pathology) is characterised by downregulation of genes such as *PVALB* (marker for myelinated inhibitory neurons) and *SV2C* (synaptic vesicle glycoprotein), consistent with the known vulnerability of PVALB+ inhibitory neurons in MS GM[19], and downregulation in oligodendrocytes of *KANK4* (Extended Data Fig. 5e), a marker for Oligo C1, an oligodendrocyte subtype with high myelin-forming potential which is lost in GML (as well as WM lesions; Extended Data Fig. 5e, bottom panel).

In WM samples, no metadata (lesions, age, sex, PMI), except donor ID, explains a large fraction of variability in each factor, highlighting that the factors capture patient-specific effects of MS (Extended Data Fig. 5c). In summary, the proportion of variance explained in the factors (Extended Data Fig. 5b,c) and the clustering of patients across factor values (Extended Data Fig. 5f,g) both suggest that patient IDs are a strong determinant of the factors identified by MOFA+. This further strengthens our conclusion from the analysis of transcriptional patterns, i.e. that given a diagnosis of MS, donor ID (patient) is a much stronger driver of variability in cell type-specific gene expression than lesion type.

Analysis of cellular heterogeneity in normal and MS brains

Next, we examined our dataset to address cellular compositional, rather than transcriptional, heterogeneity. Having defined the full extent of cellular heterogeneity in our dataset (Fig. 1b), we first examined the differences between control GM and WM. We found that four oligodendroglial subtypes (Oligos B1, B4, C1 and COP B) were significantly more abundant in WM compared to GM, which was instead enriched in other oligodendroglial subtypes COP

A1 and Oligo A2 (Fig. 3a, Extended Data Fig. 6a). GM also showed increased abundance of astrocyte subtype C, pericytes and endothelial subtype B compared to WM (Fig. 3a). Gene module analysis showed that astrocyte C is characterised by increased expression of glutamate transporter *SLC1A2* (GLIT-1), D-serine and L-serine transporter *SLC1A4* (ASCT1) and cell adhesion molecules *CADM2 ,PCDH9, NRXN1*, confirming known characteristics of GM astrocytes (Fig. 3b)[25].

Having characterised control GM and WM, we asked whether differences in cellular composition between MS WM lesions and control WM, and MS GM lesions and control GM tissue could be observed. We examined our data to robustly quantify cell compositional changes across patients and lesion types using a state-of-the art statistical method, ANCOM-BC[26]. Modelling each cell type proportion individually suggested that, just as in the transcriptional analysis, composition was strongly determined by donor for some cell types (Fig. 3c,d). We therefore fitted models to identify variability in cell type composition due to lesion type using ANCOM-BC, bootstrapping the data to account for variability due to donors (see Methods).

When comparing WM lesions, NAWM and control WM, we observed an increase in B cells, endothelial cells and pericytes in MS WM tissue, as expected in an immune-driven disease with hypoxia[27]. There was the expected increase in Microglia B, a microglial state with a more activated phenotype (Extended Data Fig. 6b), most obviously in the demyelinated lesions (AL, CAL, CIL). We found a significant reduction of oligodendrocyte progenitors (Fig. 3e, Extended Data Fig. 6c), validated by GPR17 immunohistochemistry analyses in independent samples (Fig. 3g), confirming our earlier report[6], and specific oligodendroglial populations (Oligo B2 and B3) in all lesions and NAWM, and in Oligo B1, B4, Oligo C1, COP B in most lesions and NAWM (Fig. 3e). However, as we observed in the transcriptional analysis, we did not observe lesion-specific compositional differences, but rather ‘u’/’n’-shaped profile changes (Fig. 3e).

In GM samples, we observed a strong cortical layer-specific confounding effect between GML and NAGM samples due to inclusion or exclusion of specific layers during tissue dissection: NAGM samples were significantly enriched for deeper layer neurons whereas GML lesions (which are often subpial[28]) were enriched for upper-layer neurons (Extended Data Fig. 6d). After accounting for these layer-specific effects in the ANCOM model (see Methods, Extended Data Fig. 6d,e), we observed selective loss of CUX2+ RORB+ excitatory neurons and parvalbumin inhibitory neurons in GML and NAGM/GML respectively (Fig. 3f, Extended Data Fig. 6f), confirming earlier observations[7,19]. We found significant depletion of Oligo C1 in both NAGM and GML, as seen in WM lesions (Fig. 3e). However, in contrast to WM, OPCs were not depleted in NAGM or GML. Instead, Oligo A2 was depleted in GM lesions and Oligo B1,

B2, B3 were enriched in GM lesions and/or NAGM, perhaps suggesting a different and more effective oligodendroglial regenerative response in GM compared to WM, as previously described[29]. Astrocyte B was enriched in GM lesions and NAGM, but not in MS WM, suggesting that the course of astrocytosis differs between the WM and GM regions. Microglia A and pericytes were depleted in GM, but Microglia B (which we identified as activated microglia) was not enriched, as in WM, consistent with a dissimilar microglial response in GM compared to WM[30].

These data reveal that the MS pathological micro-environments in WM and GM are distinct, as evidenced by differences in cell compositional response, and likely reflecting differences in pathological and regenerative processes. NAGM or NAWM also show differences from controls, often as a less extreme version of the corresponding lesion samples, reinforcing our knowledge that they are in fact not normal[19,31–33]. Importantly, however, although we observed differences in cellular composition between MS and control samples, we did not identify differences that were specific to any particular WM lesion type. Instead, and mirroring the findings of the transcriptional analysis, we observed a continuum of responses, with AL and CAL lesions usually showing the most extreme responses, indicating that lesion type only weakly determines cellular composition. Rather, for specific cell populations in both WM and GM lesion samples, the donor (patient) ID is the major contributor to variability in abundance (Fig. 3c,d). There is significant inter-individual variance in Oligo A2, B1, C2 and D, and Microglia A in WM samples (Fig. 3c) and Oligo B1, B2, B4 and COP A1 and subpopulations of excitatory and inhibitory neurons in GM samples (Fig. 3d). We conclude, as we did for the transcriptional analysis, that given a diagnosis of MS, donor ID (patient) is a much stronger driver of variability than lesion type in both WM and GM. Together, these two different analyses of our dataset allow the important conclusion that significant heterogeneity in the neurodegenerative and/or neuroregenerative response to MS pathology lies between patients, rather than between the different types of lesions within an individual patient.

Three distinct white matter glia patterns in MS patients characterised by oligodendrocyte subtypes

Given that much of the heterogeneity in MS pathology is determined by the individual patient, we explored stratifying the samples and donors in our study using WM MOFA+ factors, which allow identification of subgroups of patients with coherent cross-cell type transcriptional phenotypes. Hierarchical clustering based on MOFA+ WM factor scores clearly distinguishes MS patients from controls, and stratifies MS patients into distinct subgroups (Fig. 4a). With few exceptions, controls have low scores for all 5 factors, and each subgroup of MS patients

has a distinct pattern of high/low factor scores. Measures of technical quality were not different between subgroups (Extended Data Fig. 7).

To infer potential mechanisms, we examined the genes whose changes in expression were responsible for each factor (Supplementary File 6). MS patients with high factor 1 scores were characterised by upregulation of common stress-related genes across multiple cell types (e.g. *HSPA4L, HSPE1-MOB4, DNAJB1, SERPINH1* account for at least one of the top 10 MOFA genes for each cell type). In immune cells, there is high expression of immunoglobulin genes (*IGKC, IGHG3, IGHG4*), while in astrocytes, this subset of samples has reduced expression of *HES1* (Fig. 4b). Microglia also exhibited reduced expression of *CX3CR1* for Factor 1 samples (Fig. 4b). High factor 3 weighting, by contrast, is characterised by upregulation of two extracellular matrix genes, *COL19A1* and *MARCOL*, across multiple cell types (Fig. 4c). In contrast, Major Histocompatibility Complex genes such as *HLA-DPA1, HLA-DRB1 and HLA-C* were specifically reduced in microglia from Factor 3 samples (Figure 4c). Factor 5 is significant only in oligodendrocytes (Extended Data Fig. 5c), where the involvement of the genes for tyrosine kinase-like orphan receptor *ROR1*, as well as a lipid raft protein *BAALC* point to altered cellular signalling (Fig. 4d). The genes selected for input to MOFA+ are mostly not shared between cell types, suggesting that the factors identified capture coordinated tissue-level, cell type-specific responses.

As factor 5 was only significant in oligodendrocytes and due to our finding of skewed oligodendrocyte subtype heterogeneity in these data (and previous[6]), we re-examined the proportion of oligodendroglial subtypes in each WM lesion sample at the individual donors to identify whether this correlated with MOFA identified patient subgroups (Fig. 4e). Regarding individual oligodendrocyte subtype composition, there were three patterns in MS, with a striking concordance in the proportion of each oligodendroglial subtype not only in WM lesions but also in NAWM samples from each individual patient (Fig. 4e, Extended Data Fig. 8a). The first, which we term Type 1, is similar to that seen in normal WM. The second, which we term Type 2, shows high levels of Oligo D, characterised by cellular stress genes. The third, which we term Type 3, shows high proportions of the early oligodendrocyte subtype A2 and reduced levels of the later subtype Oligo C1, suggestive of an arrest at the Oligo A2 stage in the regenerative process that normally generates new oligodendrocytes following demyelination. While the Type 1 and Type 2 phenotypes were similarly represented in MS patients and controls, the Type 3 phenotype was predominantly observed in MS patients. These groupings did not associate with any patient metadata including sex, age or subtype of progressive MS, or with QC metrics (Extended Data Fig. 7) and were not present in oligodendroglia of GM (Extended Data Fig. 8b), further confirming GM and WM differences. However, we find a strong correlation between these three oligodendrocyte phenotypes (Types 1-3) and the

patient subgroupings and factors, as defined by the MOFA analysis. In Type 3 samples, either factor 1 or factor 3 predominated, while in the Type 2 samples, factor 5 was high (Fig. 4a). Therefore, these data suggest that the Type 3 phenotype may be associated with either a specific response to inflammation and stress (Factor 1 high) and/or the presence of an inhibitory extracellular matrix (Factor 3 high), with all of these observed in distinct subsets of patients as a result of intrinsic (genetic, environmental or a combination of both) differences between these patients, which need to be taken in consideration when considering precision medicine approaches.

Discussion

This single nucleus RNA sequencing study on the most extensive cohort of MS patients to date, shows that GM and WM biology in MS are fundamentally different at a molecular and cellular level. While GM changes relate to both lesion type and patient ID, WM biology is associated primarily to the diagnosis of MS and individual patient effects, and is surprisingly agnostic to classical MS lesion categories. This study takes the first step towards the stratification of progressive MS cases in terms of their molecular and cellular pathology, only made possible by the large number of samples captured both within individuals and from different individuals. Recent work has explored the trade-offs between read depth and number of individuals[34]. While greater read depth is useful for characterising lowly expressed genes and rare cell types, our study suggests that there is much to be discovered by increasing patient breadth.

Our WM results point to three fundamentally different oligodendrocyte-related phenotypes in MS, each specific for a subgroup of patients and shared by all WM lesions and NAWM in a single patient. First, one in which the regenerative process might be impaired or blocked at the Oligo A2 stage preventing the formation of sufficient oligodendrocytes for regeneration (Type 3 - stalled). Second, one in which a substantial number of oligodendrocytes show high expression of cellular stress genes (Type 2 - stressed). Finally, one in which we see no evidence for the process of oligodendrocyte developmental maturation being impaired as evidenced by the relative proportions of the oligodendrocytes subtypes being the same as in control white matter (Type 1 - standard). These phenotypes do not group with any patient metadata including age, sex, type of progressive MS, duration of disease, previous medications, post mortem delay or quality measures of the samples.

We propose these patient groupings on the basis of the presumed oligodendroglial response, but these match well to our factor analysis of gene expression across all cell types for individual patients, allowing us to interrogate different pathways across all cell types involved

in these different patient responses. This stratification into three phenotypes each with distinct transcriptional patterns as revealed by our MOFA+ analysis is important and may explain the poor response of neuroprotective/pro-regenerative therapies in progressive MS trials. We propose that current pro-remyelinating drugs that enhance oligodendrocyte formation from precursor cells may be effective only in the patient subgroup with Type 1 (standard) phenotype where remyelination is possible but simply at a low level, and not in the other subgroups where this process is blocked or with stressed oligodendrocytes. It is also conceivable that Siponimod, now approved for selected SPMS patients, has a more marked effect in the Type 2 (stressed) subgroup, through its proposed role in NRF2 signalling and antioxidant pathways[35]. However, any positive response in these trials of neuroprotective / pro-regenerative therapies may therefore be diluted out by patient heterogeneity, and effective therapies for one of the subgroups may be lost.

To stratify the living and so give future trials the greatest power to reveal effective therapies, we now need to link these post mortem phenotypes to biomarkers measurable prognostically or during disease, ideally in serum or CSF, or as potential targets for Positron Emission Tomography (PET) ligands. This will allow us to determine whether the phenotypes are stable in the same patients over time and to interrogate clinical trial datasets for effect or indeed lack of effect within post hoc stratified patient groups. In this study, we provide a resource of cell-type specific genes, identified by MOFA factors, whose expression clearly distinguishes WM and GM pathology, and subgroup phenotype to aid such future biomarker efforts. This is an essential step change for designing effective precision medicine therapeutic strategies for progressive MS - a critical unmet need.

Limitations of the study

For practical reasons, our study uses snRNAseq from post-mortem archival tissues, with the limitation that this only evaluates primarily pre-mRNA nuclear transcripts. We undersampled rare immune cell populations e.g. activated CD8 T cells, monocytes, dendritic cells, B cells and microglia inflamed in MS (MIMS)[8], which are mostly enriched at the edges of chronically inflamed WM lesions, perhaps as we focussed on the entire lesion. Nevertheless, the sharing of gene expression programs across all lesions in oligodendrocytes, astrocytes and microglia suggests the generalizability of our results. In addition, we cannot comment on subpial GM lesions adjacent to compartmentalised inflammatory meningeal infiltrates[36], as these were not in our dataset. Although we use a larger cohort of individuals than often used in single cell-omic studies, our study highlights that even larger numbers will be required to validate our findings of patient groupings based on oligodendrocyte subtype composition and factor

analysis at a transcriptional level, and the contribution of MS risk genes in such stratification. Increasing the scope of our analysis with other modalities, for instance at the proteomics, lipidomics and epigenomics level in future studies will help in the further characterization and stratification of MS patients.

Figures

Figure 1: Single cell dissection of cellular heterogeneity and differential gene expression in multiple sclerosis lesions. a Overview of the study design. **b** UMAP plot of putative cell types/states. Subclustered CNS cell types are labelled with different colours. **c** UMAP plot calculated on just oligodendroglia. **d** UMAP as in **b**, separated to show nuclei from either MS (brown) or control donors (green) in oligodendrocytes. **e** Gene modules identified within oligodendroglia clusters, annotated with top 5 genes per module. **f** PAGA applied to oligodendrocytes and OPCs / COPs across all samples (edges with weights below 0.2 not shown). **g** Number of DE genes in GM (top) and WM (bottom) respectively, in each broad cell type, at 5% FDR and minimum absolute FC of 50%, split by whether they are only DE for this combination of lesion and cell type, only DE for this cell type (i.e. is DE in multiple lesion types for this cell type), or are also DE in another cell type (in principle such genes could be DE only in one lesion type, in practice there are very few such genes). opc_cop = OPCs and COPs; oligo = oligodendrocytes; astro = astrocytes; micro = microglia; excit = excitatory neurons; inhib = inhibitory neurons; endo = endothelial cells; peri = pericytes; immune = T and B cells. **h** Dotplots of Hallmark gene set enrichment results for GM (top) and WM (bottom). **i** Differentially expressed genes related to glutamate signalling, glucose homeostasis, ion channels and oxidative phosphorylation in excitatory neurons in GM lesions. Asterisks show FDR of this gene, using Benjamini-Hochberg correction across all genes tested for this combination of cell type and model coefficient; symbols indicate FDR, *** < *0.001* < ** < *0.01* < * < *0.05* < . < *0.1* < *1*. **j** Differential expression of interferon alpha and gamma genes in oligodendroglial cells in WM lesions. FDR annotated as in **i**. Asterisks in gene name indicate whether this gene was significantly associated with MS genetic risk from GWAS (MAGMA gene-level FDR <5%). **k** Enrichment of cell type-specific differentially expressed genes in MS WM and GM lesions with MS GWAS risk genes. The mean strength of association (−log10P) of MAGMA is shown and the bar colour indicates whether up or downregulated genes are significantly enriched in MS risk genes.

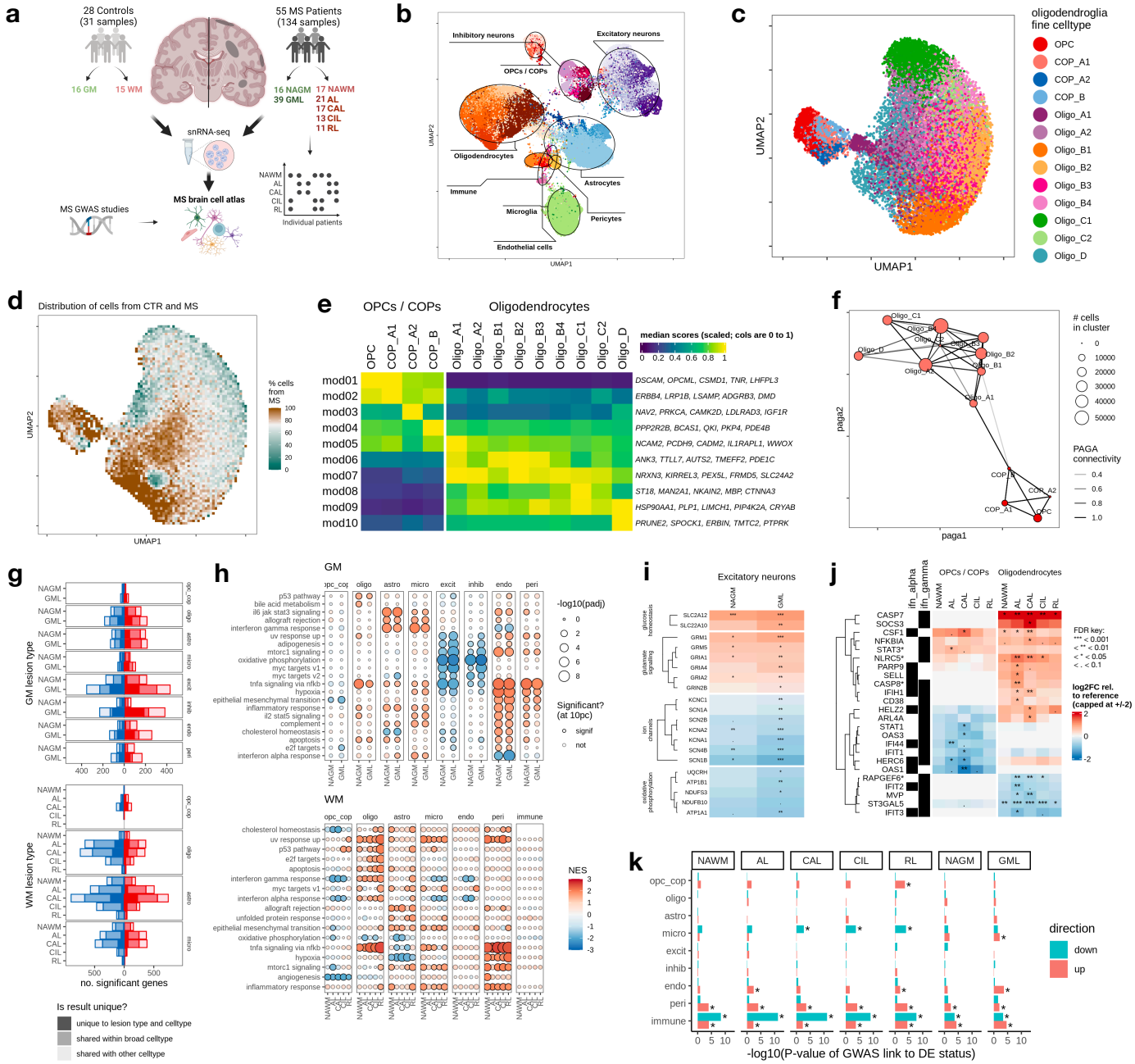


Figure 2: Patterns and determinants of cell-type specific gene expression profiles in WM and GM lesions. a Clustering of WM gene expression fold change profiles over lesion types for each broad cell type (opc_cop = OPCs and COPs; oligo = oligodendrocytes; astro = astrocytes; micro = microglia). Restricted to genes where at least one lesion type has FDR < 5%, hierarchical clustering with cut distance set to log(4), and clusters with fewer than 5 genes not shown. Figure in brackets shows number of genes in the cluster. **b** Pseudobulk expression heatmap of genes showing either MS or donor variability in broad cell types in WM, samples ordered by lesion type. **c** Pseudobulk expression heatmap as in **b**, row order on basis of hierarchical clustering. Shows that differences between MS donors are not explained by lesion type, sex, or type of MS, but all samples from one donor cluster together, suggesting a strong donor effect. **d** Pseudobulk expression heatmap of genes showing either MS or donor variability in broad cell types in GM, samples ordered by lesion type. **e** Pseudobulk expression heatmap as in **d**, row order based on hierarchical clustering. Column order in **b-e** determined by the first principal component of the gene matrix for each cell type.

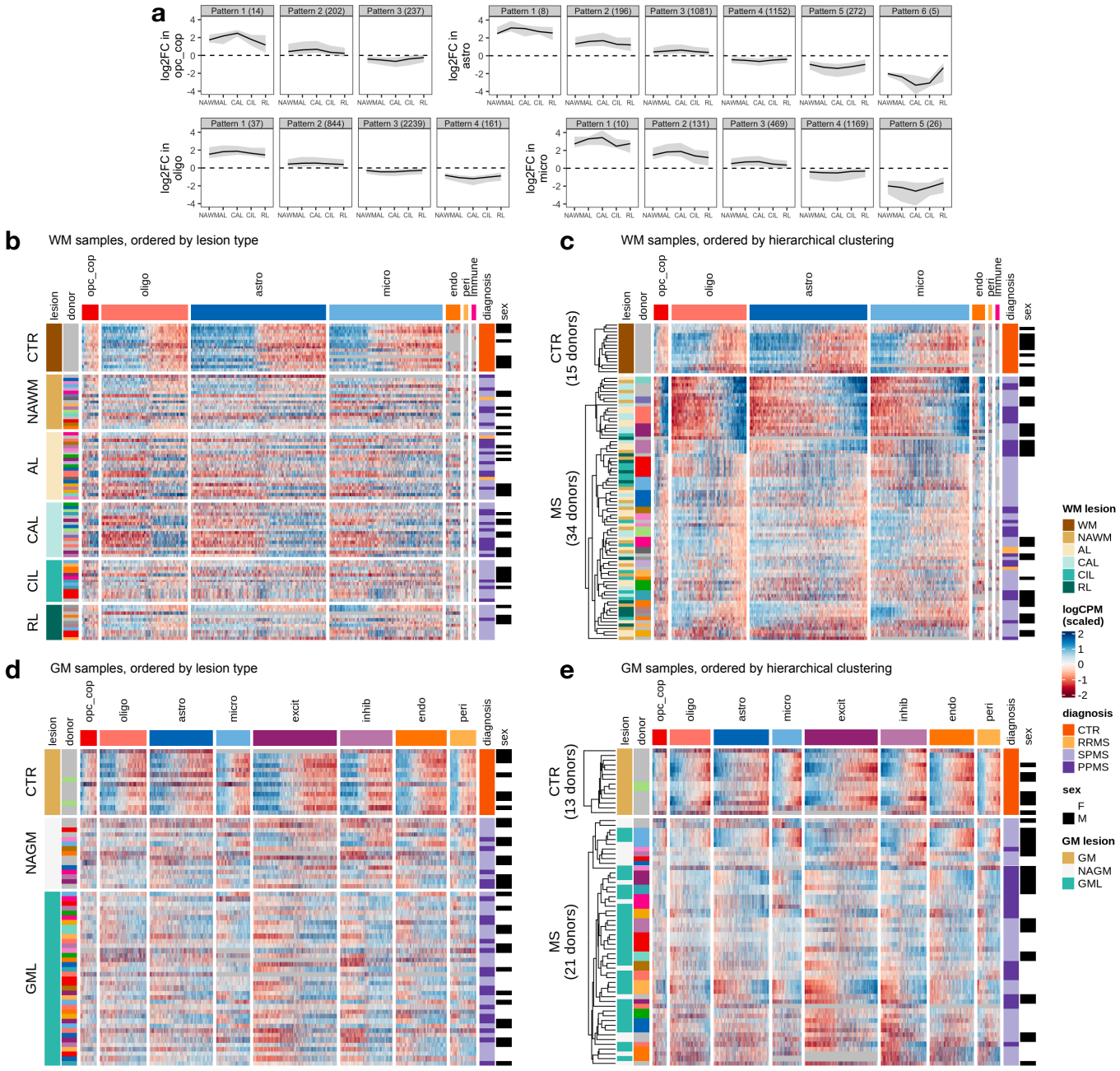


Figure 3: Differential abundance analysis of cell types in MS. a Differential composition by proportions of fine glial cell types for control GM and WM. Neuronal cell types excluded. Negative binomial model fit to absolute numbers for each cell type with log number of cells in sample as offset (see Methods). FDR calculated across all cell types; cell types significant at < 5% FDR. **b** Gene modules identified within astrocyte clusters, annotated with top 5 genes per module. **c,d** Contribution to variability in cell type abundances explained by lesion plus patient in WM (**c**), and GM (**d**) (see Methods). In **d**, four neuronal layer PCs are included as confounders. FDR is Benjamini Hochberg-corrected p-values from likelihood ratio tests of nested models. Axes show evidence that including lesion type (y axis) or donor (x axis) significantly improve the fit of a model explaining cell type proportion of a sample. Dashed lines show 5% FDR. **e** Differential abundance of WM lesion samples against control WM, as calculated by bootstrapped ANCOM-BC (see Methods and Supplementary Note). Dashed line at 0 corresponds to no difference between control and lesion. Model fitted is *count ~ lesion_type + sex + age_scale + pmi_cat* (where *age_scale* is age at death, normalized to have SD = 0.5[37], and *pmi_cat* is post-mortem interval split into three categories; see Methods). Point corresponds to median bootstrapped log2FC effect estimated by ANCOM-BC; coloured range is 80% bootstrapped confidence interval, grey range is 95% CI. Points are filled when the 95% CI excludes zero; otherwise empty. **f** Differential abundance of GM MS samples against control GM, as calculated by bootstrapped ANCOM-BC; otherwise as for **e**. Model accounts for sample layer position, by using formula *count ~ lesion_type + sex + age_scale + pmi_cat + layer_PC1 + layer_PC2 + layer_PC3 + layer_PC4* (see Methods). **g** Immunohistochemistry validation of number of GPR17-expressing cells in independent MS WM samples. Upper panel shows the number of cells with non-zero expression of GPR17 in each WM snRNAseq sample included in this study. Lower panel shows mean GPR17+ cells per mm2 (total calculated over 10 randomised fields) in an independent cohort of MS WM samples (see Methods). In both panels: horizontal line denotes median; colour denotes donor, distinct sets of donors used for snRNAseq and immunohistochemistry.

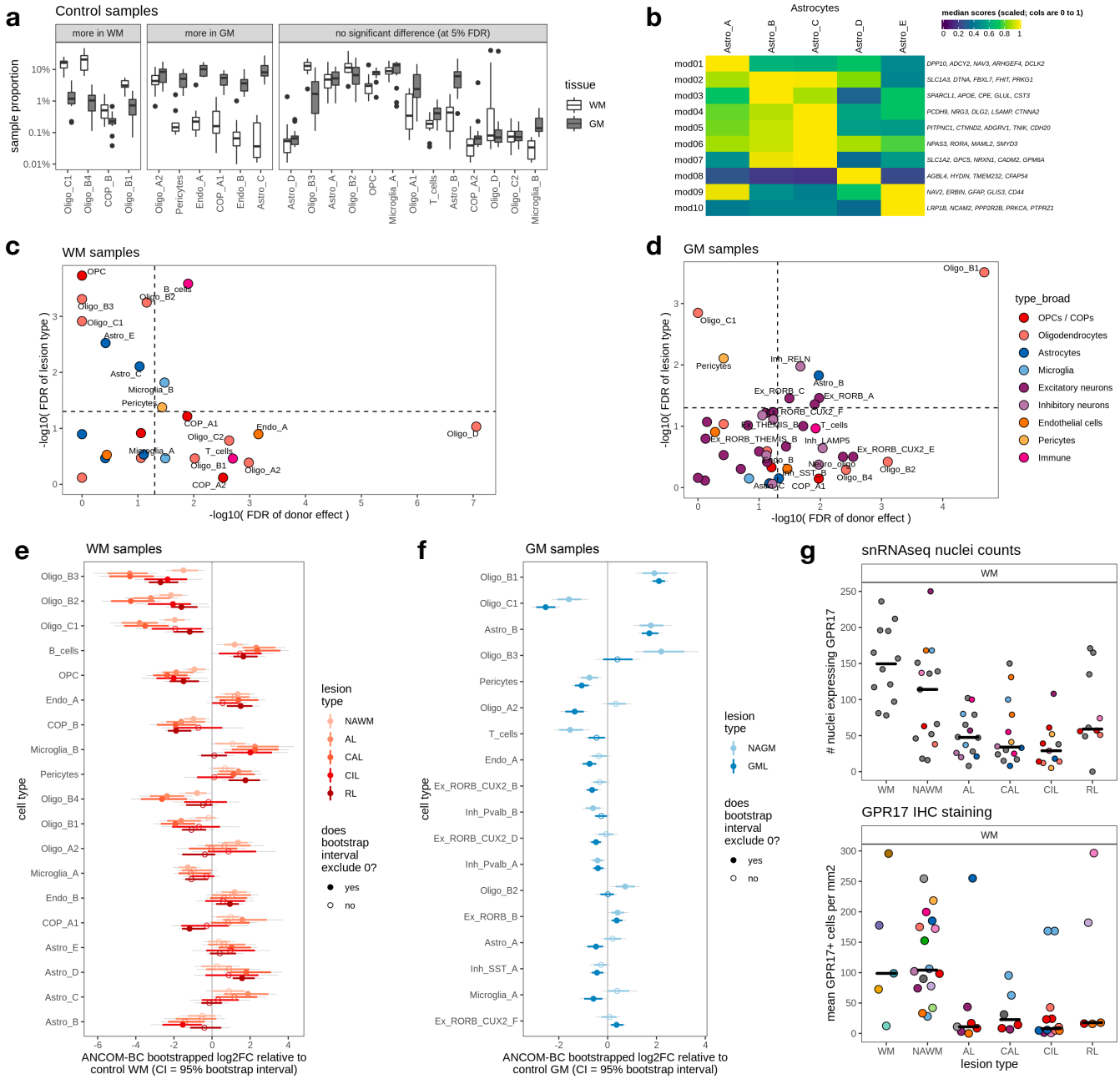
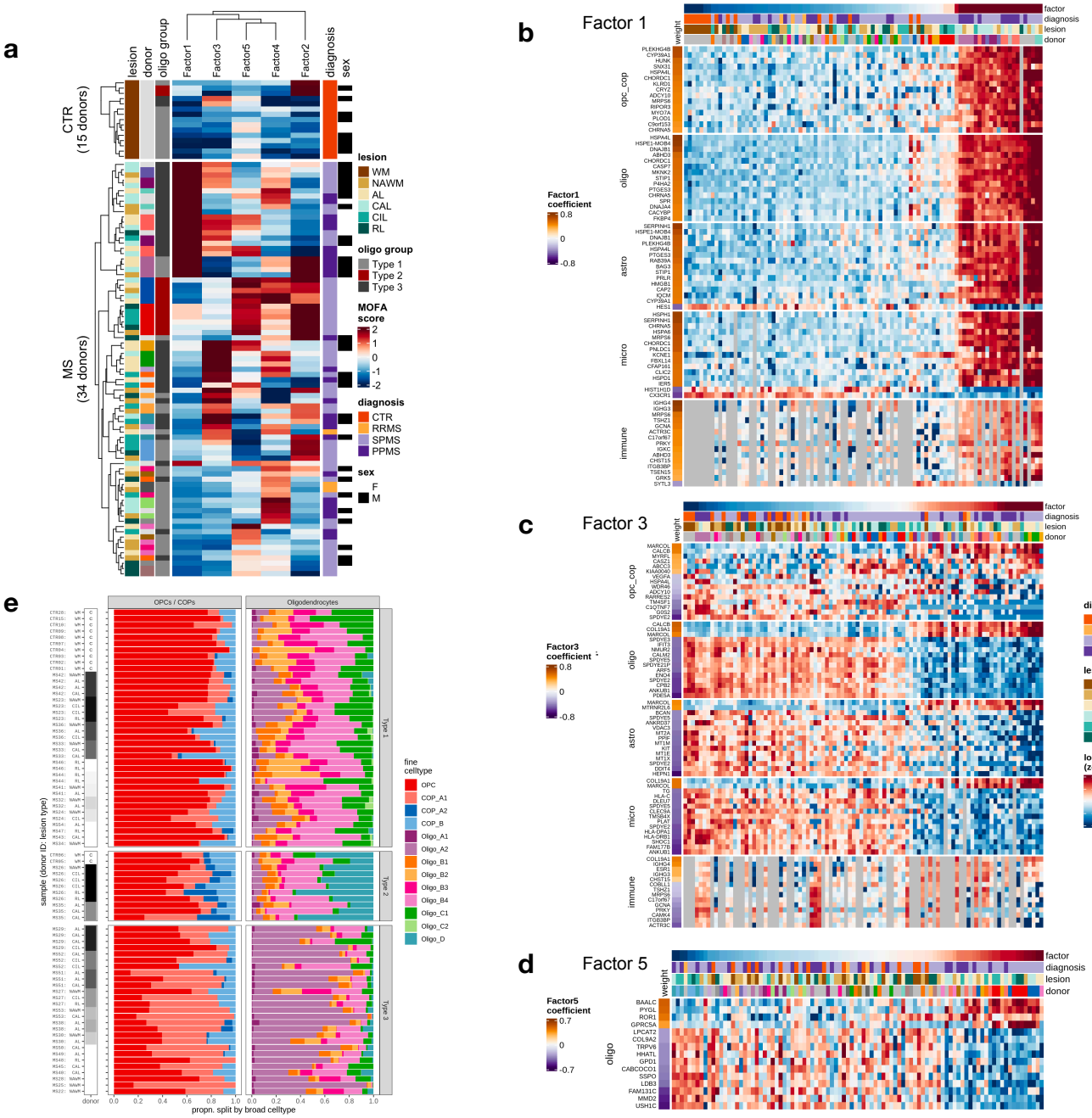


Figure 4: WM patient stratification. a Patient stratification via MOFA+ factors. Rows are samples, annotated by metadata and oligo groupings. Columns are MOFA+ factors, with signs flipped to positively correlate with MS status. **b** 15 genes with largest absolute Factor 1 weights for each cell type where >= 10% variance is explained by Factor 1. In this plot, we exclude lincRNAs and show only protein coding, IG and TCR genes; the genes used as input to MOFA include lincRNAs. Columns are samples, ordered in increasing order of Factor 1 score. Rows are genes, split by broad cell type. Heatmap colours are log CPM of pseudobulk expression, z-scored within each row; grey indicates there were less than 10 cells of this cell type in the sample. Column annotations show lesion type, diagnosis and MOFA factor score. Row annotations show MOFA factor weight of each gene. **c** 15 genes with largest absolute Factor 3 weights for each cell type where >= 5% variance is explained by Factor 3. Otherwise as for **b**. **d. 15** genes with largest absolute Factor 5 weights for each cell type where >= 5% variance is explained by Factor 5. Otherwise as for **b,c**. **e** Bar chart of proportions of WM oligodendroglia grouped by individual donor, including all samples from that donor. Each row is annotated as *anonymised donor id: lesion type*. C = control white matter; NAWM= normal appearing white matter; AL=active lesion; CAL= chronic active lesion; CIL=chronic active lesion, RL= remyelinating lesion. In this panel, WM samples with unusually high proportions of neurons were excluded (see Methods).



Methods

Sample preparation and single nuclei RNA sequencing

Brain tissue samples, ethical compliance and clinical information

Human tissue samples were obtained from the Netherlands Brain Bank (NBB), the MS UK tissue bank (UKTB) and the Edinburgh Brain Bank (EBB) via donor schemes with full ethical approval from respective brain banks (METc/2009/148 from Medical Ethical Committee of the Amsterdam UMC, MREC/02/2/39 from UK Ethics Committee), and individual material transfer agreements between Roche and NBB, UKTB and EBB. We have complied with all relevant ethical regulations regarding the use of human postmortem tissue samples. We examined a total of 114 (84 MS and 30 controls) snap frozen brain tissue blocks obtained at autopsies from 55 MS patients and 30 controls. MS patients and controls were matched for age and sex. For detailed donor information see Supplementary File 1.

Brain tissue characterization

Snap frozen tissue blocks from donors with GM lesions were provided by UKTB to Roche. Subpial GM lesions were determined using MBP (abcam-ab6263) by neuropathologists at Roche and confirmed by independent experts (Anna Williams, Roberta Magliozzi). Pathological staging of WM lesions from EBB and NBB donor samples was done at the respective brain banks. In the WM, de- and remyelinated lesions were identified by Luxol Fast Blue (LFB) staining, and immunohistochemistry for proteolipid protein, and demyelinated lesions were grouped into active, chronic active and chronic inactive lesions with Oil red O and/or HLA-DR staining to determine microglial activity[38]. Brain tissue specimens from the respective WM regions were shipped on dry ice to Roche and directly processed.

Nuclei isolation and single nuclei RNA sequencing

Nuclei were isolated from fresh-frozen 10µm sections, using Nuclei Pure Prep Nuclei Isolation Kit (Sigma Aldrich) with the following modifications. The regions of interest were macro-dissected with a scalpel blade, lysed in Nuclei Pure Lysis Solution with 0.1% Triton X, 1mM DTT and 0.4U/ul SUPERase-In™ RNase Inhibitor (ThermoFisher Scientific) freshly added before use, and homogenised with the help first of a 23G and then of a 29G syringe. Cold 1.8M Sucrose Cushion Solution, prepared immediately before use with the addition of 1mM

DTT and 0.4U/ul RNase Inhibitor, was added to the suspensions before they were filtered through a 30µm strainer. The lysates were then carefully layered on top of 1.8M Sucrose Cushion Solution. Samples were centrifuged for 45min at 16000xg at 4°C. Pellets were resuspended in Nuclei Storage Buffer with 0.4U/ul RNase Inhibitor, transferred in new Eppendorf tubes and centrifuged for 5 min at 500xg at 4°C. Pellets were again re-suspended in Nuclei Storage Buffer with 0.4U/ul RNase Inhibitor, and centrifuged for 5 minutes at 500xg at 4°C. Finally, purified nuclei were re-suspended in Nuclei Storage Buffer with 0.4U/ul RNase Inhibitor, stained with trypan blue and counted using an automated cell counter (Countess II, Life technologies). A total of 12,000 estimated nuclei from each randomised sample was loaded on the 10x Single Cell Next GEM G Chip. cDNA libraries have been prepared using the Chromium Single Cell 3' Library and Gel Bead v3.3 kit according to the manufacturer's instructions. cDNA libraries were sequenced using Illumina NovaSeq 6000 System and NovaSeq 6000 S2 Reagent Kit v1.5 (100 cycles), aiming at a sequencing depth of minimum 30K reads/nucleus.

snRNAseq data processing and quality control

All samples were processed with CellRanger (v3.1.0), using the GRCh38 reference human genome and the ensembl Homo_sapiens GRCh38.96 reference annotation (modified to count intronic reads). Gene expression quantifications for each nucleus were obtained from the 'filtered_feature_bc_matrix' CellRanger (v3.1.0) output folder. In addition, we used velocyto (v0.17.17) on the CellRanger output to quantify intronic and exonic reads. We identified putative doublets using `scDblFinder` (version 1.4.0), applied to each sample separately (*multiSampleMode* = *"split"*), with all other parameters default[39].

After removing doublets, we did quality control with SampleQC (version 0.4.5)[40]. We first removed nuclei and samples with insufficient data to be worth including, requiring nuclei to have a minimum of 300 expressed genes, 500 UMI counts, and samples to have at least 500 cells remaining after this first filtering. As inputs to SampleQC, we used the metrics *log10* transformed library size (*log10_counts*), inverse logistic-transformed mitochondrial proportion (*logit_mito*), and the log2-transformed ratio of spliced to unspliced reads (*splice_ratio*). We ran SampleQC with default parameters, grouping the 173 samples into 9 groups with similar QC metric distributions. We fit separate models to each of these groups, setting the number of clusters assumed for each group by inspection, with these numbers ranging between 1 and 4. In many samples, we observed a cluster of cells that had both high splice ratios and high mitochondrial proportions: in this cluster ~10% of reads were unspliced; for most other clusters, 60%-80% of reads were unspliced. We therefore excluded any cells assigned to a

cluster with mean log2 splice ratio higher than 3 (equivalent to ~88% spliced reads). After excluding outliers, we again excluded any samples with fewer than 500 cells remaining. The full SampleQC report is included as Supplementary File 2. This resulted in 740,023 nuclei across 165 samples passing QC.

Sample swap checks via genotyping

We genotyped all samples included in this study using the GSAv3 illumina CHIP. Genotypes were imputed using the Haplotype Reference Consortium (HRC) reference panel (version r1.1)[41] and lifted over to GRCh38. Genotype processing and quality control was performed using Plink v1.9[42]. SNPs with imputation score <0.4 or with missingness greater than 5% were excluded. We used MBV[43] to identify sample swaps. Briefly, MBV takes as input a VCF file containing the genotype data of the samples, as well as bam files containing the mapped single nuclei sequencing reads. MBV then reports the proportion of heterozygous and homozygous genotypes (for each individual in the VCF file) for which both alleles are captured by the sequencing reads in all bam files. Correct samples can then be identified as they should have a high proportion of concordant heterozygous and homozygous sites between the genotype data and the mapped sequencing reads. We identified and corrected 3 sample swaps; one sample was excluded because it could not be matched to a genotype.

Data integration and clustering

Data integration

Data integration was done with Conos (version 1.4.0)[11], applied to the highly variable genes in each sample, with highly variable genes identified by sctransform (version 0.3.2)[44]. Clustering was done with the function findCommunities at resolution 8, and any cluster with fewer than 100 cells in total across the whole dataset was excluded, resulting in 55 clusters. Success of the integration was checked by calculating the entropy of the cell counts across all samples for each cluster; this analysis was also done at the donor level, and for various other metadata variables (Extended Data Fig. 2d).

Broad cell types were assigned to each cluster on the basis of known marker genes: *PLP1, MAG, MOG, OPALIN* (Oligodendrocytes), *PDGFRA, BCAN* (OPCs / COPs), *FGFR3, GFAP, SLC14A1, AQP4* (Astrocytes), *P2RY12, SPP1, CSF1R, IRF8 (Microglia), SLC17A7, FEZF2, RORB* (Excitatory neurons), *GAD1, ADARB2, LHX6* (Inhibitory neurons), *CLDN5, FLT1, EPAS1* (Endothelial cells), *EPS8, LAMA2* (Pericytes), *IGHG1* (B cells) and *IL7R (*T cells). The

log normalised expression of each gene was calculated for each cluster, and these values scaled to 0 to 1 over all clusters. For each cluster, the broad cell type with the highest scaled expression averaged over the known marker genes was selected as the label.

Subclustering

For some clusters, we had strong biological expectations that further subclustering was necessary: cluster cns24 (possibly a mixture of neurons and oligodendrocytes), cluster cns28 (OPCs and COPs) and cluster cns55 (Immune cells). These clusters were further subclustered using the batch-corrected graph from Conos to obtain a subgraph for each individual cluster, then by rerunning findCommunities (with resolutions 0.5, 0.5 and 1 for clusters 24, 28 and 55 respectively). To check the identities of these clusters, we plotted expression of the broad cell type marker genes listed above, a selected list of immune genes, and markers selected by running findMarkers from the scran package (version 1.18.7) within each set of subclusters[45].

Merging non-distinct clusters with SCCAF

We merged insufficiently distinct subclusters using SCCAF (version 0.0.10)[12]. To decide whether a pair of clusters should be merged, SCCAF uses the maximum probability of misclassification in either direction. We found that this metric was too aggressive, resulting in merging of clusters that marker expression indicated should be kept separate. We therefore amended the SCCAF procedure to use the misclassification error calculated over all cells of both clusters in the pair. Using this amended approach, we merged clusters with a probability > 10% of being confused by the SCCAF classifier. In addition to this, based on marker expression and biological expectations, we defined a custom list of clusters that would either be merged or kept unmerged: two Ex_RORB_CUX2_A subclusters were merged (cns03 and cns05); three T cell subclusters were merged (cns55.1, cns55.4, cns55.5); and each of Neuro_oligo (cns24.2), COP_B (cns28.3), COP_A2 (cns28.4) and B cells (cns55.3) were kept separate. This resulted in 50 merged final clusters.

We used an independent data integration and clustering method (Seurat v4[16] + Harmony[17]) using QC passed samples and found high concordance between clusters identified by both methods (Extended Data Fig. 2c), thereby validating our approach.

Filtering of genes contaminated by ambient RNA

We used the maximumAmbience function in the package DropletUtils version 1.10.3[46]. In each sample, we used droplets with < 100 reads and calculated the proportional expression using

Good-Turing frequency estimation to obtain a profile, with other parameters set as defaults. Genes with mean ambient proportion greater than 10% across samples with non-zero expression were excluded, separately for each cell type. Such genes were excluded in cell type annotation, differential expression and MOFA+ analysis.

Cell type annotation

We assigned cell type identity to clusters based on known cell lineage markers, as well as by comparison with previously published and reported snRNA-seq studies[6,47]. We identified the top differentially expressed genes for each subcluster of a broad cell type by running findMarkers from the scran package (version 1.18.7) within each set of subclusters. Furthermore, we ran orthogonal non-negative matrix factorization (oNMF, implemented in PopAlign[48] version 0.1) to identify distinct gene modules and individual genes for each cluster as potential cluster marker genes. We ran oNMF separately for each of the following groups of clusters: OPCs / COPs and oligodendrocytes; microglia and immune; excitatory neurons; inhibitory neurons; astrocytes; endothelial cells and pericytes. For each fine cell type within these groupings, we selected up to 2,000 nuclei at random (where there were fewer than 2,000 nuclei in the cluster, we used all nuclei). Any genes with contamination estimated to be greater than 10% were set to zero, and restricted to genes with the *protein_coding* biotype. We then used the SCTransform function in the Seurat package to identify the top 2,000 highly variable genes. We used the onmf function in popalign using these highly variable genes, and otherwise default parameter values.

Differential abundance of cell types in MS lesions and control samples

We first checked for samples with sample sizes that were much smaller than for other samples, to exclude samples where abundances might be very noisy. We excluded samples with log sample sizes 2*MAD (median absolute deviation) less than the median log sample size, separately for WM and GM; this excluded zero WM samples and two GM samples. We also checked for samples with unusual proportions of neuronal cells relative to other WM or GM samples. White matter samples with neuronal proportion at least 2*MAD (median absolute deviation) higher than the median neuronal proportion for all WM samples were excluded; grey matter samples with at least 2*MAD neuronal proportion fewer than the GM median neuronal proportion were excluded. This excluded 18 out of 94 WM samples and 1 out of 71 GM samples.

To test whether abundances of fine cell types across samples were affected by lesion type and donor ID (**Fig. 3b,c**), we used likelihood ratio tests applied to models including lesion type and donor id. Briefly, we fit a series of nested models for each fine cell type: full (*counts ~ lesion_type + sex + age_scale + pmi_cat + (1 | donor_id)*), fixed (*counts ~ lesion_type + sex + age_scale + pmi_cat*), covariates only (*counts ~ sex + age_scale + pmi_cat*) and null (*counts ~ 1*). We used the R package *glmmTMB* (v1.1.2)[18] to fit a negative binomial distribution to the raw counts for each cell type (see Supplementary Note for explanation of use of raw counts rather than proportions). We used the function *anova* to perform likelihood ratio tests of the following nested sequence of models: full; fixed; covariates only; null. This gives a p-value indicating whether the more complex model improved the fit more than would be expected by chance. We adjusted the p-values across these tests using the Benjamini-Hochberg procedure, across all cell types and models together.

To test for differential abundance in fine cell type due to lesion type, we used ANCOM-BC version 1.3.2[26]. The likelihood ratio test analysis above indicated that donor ID needed to be taken into account, however this version of ANCOM-BC does not allow random effects. To factor out donor effects, we therefore did a bootstrapped analysis of abundance: each bootstrap took one random sample per donor, and ran ANCOM-BC on each bootstrapped sample (e.g. in WM, there were 76 samples across 42 donors, so each bootstrap comprised 42 samples). We summarised the results of 20k bootstraps by taking the median, 80% and 95% confidence intervals of the inferred coefficients for each fine cell type (20k samples is sufficient to properly estimate tail probabilities for 95% CIs; cf [49]).

To test differential abundance in WM samples, we additionally excluded all neuronal cell types, as these should not be present in WM. We fit ANCOM-BC with the formula *~ lesion_type + sex + age_scale + pmi_cat*, where *age_scale* is patient age, scaled to have SD = 0.5 across all patients in the dataset[37], and *pmi_cat* is post-mortem interval, split into three categories (under 1 hour, between 1 and 12 hours, and more than 12 hours).

To test differential abundance in GM samples, we first fit the data with a similar formula: *~ lesion_type + sex + age_scale + pmi_cat2* (here, *lesion_type* includes ctrGM, NAGM and GML; *pmi_cat2* has only two categories to reflect the values observed in GM, between 1 and 12 hours, and more than 12 hours).

Using this model produced results that conflicted with known biology, identifying multiple neurons as increasing in abundance in GM lesions relative to GM controls. Analysing differences between neuronal proportions between ctrGM, NAGM and GML, we observed that GML samples were enriched in L1/L2/L3 neurons, and those from NAGM samples were enriched in L5/L6 neurons (ctrGM samples had intermediate proportions). This indicated that

GML samples were taken from higher in the brain, and the matched NAGM were taken from beneath them (although the experimenters had made efforts to take all samples from the same depth).

To identify layer effects for each sample, we calculated principal components (PCs) reflecting neuronal layer distributions in normal tissue. We applied PCA to the centred log ratios of the neuronal cell types in the control GM samples. We then identified principal components that could be relevant to layers (by filtering on both the absolute Spearman rank correlation between the PC loading and the layer numbers of neurons known to be layer-specific, thresholding at minimum 0.2 correlation) and which explained at least 1% of variance. This identified seven PCs that could contain layer information (Extended Data Fig. 6e). This analysis was performed in control GM samples only; we then calculated CLRs for all samples, and projected them into the selected PCs, using the loadings derived from the control samples.

We then repeated the bootstrapped ANCOM-BC analysis, including layer PCs as covariates to factor out layer effects. We used the formula ~ *lesion_type* + *sex* + *age_scale* + *pmi_cat2* + *layerPC1* + … + *layerPCn*, i.e. we repeated the analysis using the first *n* layer PCs identified above, including from 1 up to 7 PCs. We found that including the first 4 layer PCs gave results that fitted well with expected biology, i.e. that almost no neuronal types were found to increase in abundance in either NAGM or GML relative to control GM, and PVALB+ neurons decreased in abundance (Extended Data Fig. 6g). We included 4 layer PCs, however there is little difference in the results for including between 3 and 7 layer PCs.

Differential expression analysis using generalised linear mixed models

To identify genes differentially expressed in MS WM and MS GM samples compared to respective control samples per cell type, we performed a differential expression analysis on pseudo-bulk data, i.e. analysis at the level of the transcript totals across all cells of a given type in each sample. Pseudobulk approaches are known to offer a good compromise between sensitivity and run time constraints[50,51] (see Supplementary Note for the details of our analysis using different pseudobulk approaches and identification of strong patient effects). Inspection of gene expression at the donor level indicated that our model would need to include donor effects. We therefore used glmmTMB[18] with a negative binomial model, and donor_id as a random intercept. The formula for WM was *counts* ~ *lesion_type* + *sex* + *age_scale* + *pmi_cat* + *(1* | *donor_id)*, where *pmi_cat* is post-mortem interval, split into three categories (under 1

hour, between 1 and 12 hours, and more than 12 hours, and *age_scale* is patient age, normalised to have *SD = 0.5*[37].

In the GM analysis, we accounted for layer effects by including 4 layer PCs as described in the ANCOM-BC analysis. The formula for GM was therefore *lesion_type + sex + age_scale + pmi_cat2 + ctrl_PC01 + ctrl_PC02 + ctrl_PC03 + ctrl_PC04 + (1 | donor_id);* to reflect values observed in GM samples, *pmi_cat2* has only two categories (between 1 and 12 hours, and more than 12 hours). We included an offset of *log(lib.size) - log(1e6)*, so that the reported coefficients correspond to log counts per million (logCPM). Genes with absolute log2 fold change in expression of at least *log2(1.5)* and an FDR-corrected $P < 0.05$ were selected as differentially expressed. FDRs were calculated at the level of combination of cell type and model coefficient.

To quantify the extent of donor effects, for each gene we also used glmmTMB to fit three simpler models: with fixed effects only (*counts ~ lesion_type + sex + age_scale + pmi_cat*), with covariates only (*counts ~ sex + age_scale + pmi_cat*) and a null model (*counts ~ 1*). We then used the anova function to perform likelihood ratio tests for this nested sequence of models; we applied a Benjamini-Hochberg correction across all genes and LRTs, separately within each cell type.

Gene set enrichment analysis of differentially expressed genes

FGSEA[52] was used to perform statistical enrichment tests of differentially expressed genes in each cell type (broad and fine) from each comparison in WM and GM samples. All genes expressed in a given cluster were used as a background list, and GO-term analysis for enriched biological processes and hallmark genes from MSigDB[53] was performed. FDR correction was calculated within each combination of cell type, model coefficient and pathway collection. Processes with an FDR-corrected $P < 0.1$ were considered and their normalised enrichment scores (NES) plotted as a dotplot using ggplot2[54] R-based libraries.

MAGMA analysis of differentially expressed genes

We used MAGMA v1.08[21] to test whether the top 1000 most differentially expressed genes (ordered by p value) in each cell type and lesion type were enriched in MS genetic associations[22] (by direction of effect). Briefly, MAGMA computes gene-level p values from GWAS summary statistics by aggregating p values from all SNPs around each gene (here set as 35kb upstream to 10kb downstream of each gene) accounting for linkage disequilibrium (here using the European reference panel from phase 3 of the 1000 Genomes Project[55]).

These gene-level p values are then transformed to z-scores. A linear regression is then performed to test whether genes in an input gene set have higher z-scores than the rest of the genome (one sided test). The linear model from MAGMA accounts for the following covariates: gene size, gene density, mean sample size for tested SNPs per gene, the inverse of the minor allele counts per gene and the log of these metrics. Before selecting the top 1000 most differentially expressed genes, we only retained genes with a gene-level p value in MAGMA (i.e. only protein-coding genes with tested GWAS SNPs in their vicinity).

Assessment of cluster connectivity with PAGA

To characterise connectivity between clusters, we used PAGA[14] as implemented in scanpy version 1.8.2[56]. As input, we used the nearest neighbour graph constructed by conos, restricted to just cells with oligodendrocyte or OPC / COP labels. We ran PAGA clustering and layout embedding using fine cell types as the group variable, and otherwise used defaults.

Patient stratification using MOFA+

We used MOFA+ to identify factors explaining the variability across the samples (implemented in the R package MOFA2, v1.4.0)[24]. MOFA+ was developed for data with multiple modalities measured from the same samples. In this study, we took the different *cell types* to be the different modalities. This allows us to identify responses that are coherent across samples, across multiple cell types simultaneously, and which may have cell type-specific responses. MOFA+ does this by finding factors that seek to explain the variability in the input data (intuitively similar to PCA) across samples, and which correspond to coordinated tissue-level responses, even though the genes identified for each cell type are distinct. For example, a factor corresponding to remyelination might involve myelination genes in oligodendrocytes, debris-clearance genes in microglia, and metabolic support genes in astrocytes. These factors can then suggest possible ways to stratify patients.

We first identified genes with relevant variation for each cell type, based on the negative binomial models fitted to each gene in each cell type. We identified genes with either an MS effect, or a donor effect (or both). Genes with an *MS effect* were defined as those where at least one lesion type had both an FDR < 1% and an absolute log2 fold change of 1 for WM, or log2(1.5) for GM (we observed lower effect sizes in GM, and therefore used a more relaxed threshold). Genes with a *donor effect* were those where the likelihood ratio test of including the donor effect had an FDR < 1%, and the standard deviation of the donor random intercepts was at least *log(2)* for WM, and *log(1.5)* for GM. These thresholds are arbitrary but we have

found the factors identified by MOFA+ to be robust to variations on these thresholds. In WM, this resulted in the selection of: 97 genes for OPCs + COPs, 507 for oligodendrocytes, 794 for astrocytes, 667 for microglia, 86 for endothelial cells, 27 for pericytes and 27 for immune cells. In GM, we selected: 187 genes for OPCs + COPs, 531 for oligodendrocytes, 721 for astrocytes, 385 for microglia, 952 for excitatory neurons, 597 for inhibitory neurons, 582 for endothelial cells and 298 for pericytes.

We then calculated normalised expression for the selected genes in each cell type as input to MOFA+. We first excluded any samples with fewer than 10 cells observed for that cell type (this means that there may be missing data for some cell types in some samples). We also excluded any samples with log library sizes 3*MAD (median absolute deviation) less than the median log library size for that cell type. For the remaining samples, we calculated the *log(CPM + 1)* of the pseudobulk values, calculating CPMs with library sizes via the effectiveLibSizes function in edgeR[57]. To remove any possible layer effects in GM samples, we fit a linear model using the first four layer PCs as covariates, i.e. *logCPM ~ layerPC1 + layerPC2 + layerPC3 + layerPC4*, and used the residuals of this model as values for input to MOFA+. To ensure each gene contributed equally to the model, we then z-scored all resulting values within each combination of gene and cell type.

For each of WM and GM, we then fit MOFA+ to this data, using 5 factors. As we are interested in an unbiased characterisation of the heterogeneity of the data, we did not use the group variable in MOFA+; otherwise we used the default parameters. In both WM and GM, we found 5 factors which explained at least 5% of variance for some cell type.

To calculate gene set enrichment of the genes in MOFA+ factors, we ranked the genes for a given cell type in descending order of the factor weight, and used the function fgseaMultilevel from the FGSEA package[52], using a minimum set size of 5 genes and otherwise the default parameters.

Immunohistochemistry and analysis

FFPE tissue sections from EBB were used for immunohistochemistry staining for GPR17. These comprised 26 blocks in total (5 control WM, 16 NAWM, 7 AL, 6 CAL, 12 CIL, 5 RL) from 20 donors; donors were distinct from those used for snRNAseq.

FFPE sections (4 µm) were deparaffinized in decreasing concentrations of ethanol, and antigen retrieval was performed in antigen unmasking solution (Vector Laboratories, H-3300) for 10 min in the microwave. Sections were incubated with autofluorescence eliminator reagent (Millipore, 2160) for 1 min and washed with TBS 0.001% Triton-X (wash buffer).

Endogenous peroxidases were quenched with 3% H_2O_2 for 15 min at room temperature (RT), washed in wash buffer and blocked for 30 min at room temperature with PBS 0.5% Triton-X (TBS-T), 10% HIHS (blocking buffer). Primary antibody incubation was performed overnight at 4 °C in blocking buffer. Sections were washed and incubated for 2hrs at RT with HRP-labelled secondary antibodies. Fluorophore reaction was performed using OPAL 570 and OPAL 650 reaction kits for 10 min at RT (Akoya Biosciences, FP1488001KT and FP1496001KT respectively, 1:500). Sections were counterstained using Hoechst (Thermo Fisher, 62249; 1:1,000), washed and mounted.

The following primary antibodies were used: mouse anti-CNP (Atlas, AMAb91072, 1:1,000), rabbit anti-human GPR17 (Cayman Chemical, 10136, 1:100), MBP (abcam-ab6263), PLP ([]). The following secondary antibodies were used: Vector Laboratories, rabbit-HRP IgG (MP-7401, Vector laboratories), mouse-HRP IgG (MP-7402, Vector laboratories).

For quantifications of GPR17 cell numbers, sections were co-labelled with GPR17 and CNP which was used to define demyelinated lesions. Sections were scanned using a VectraPolaris slide scanner and processed using Qupath[58] and Fiji[59] imaging software. Within individual lesions, several regions of interest were selected randomly. These regions of interest were randomised using the Fiji randomization plugin and quantified completely blinded mixing samples from all conditions, regions and lesions.

Statistical analysis

No statistical methods were used to predetermine sample sizes, but our sample size is eight times larger than those reported in previous snRNAseq MS publications (Jäkel et al.[6], Schirmer et al.[7], Absinta et al.[8]). Statistical analyses and graphical visualisations were performed using open-source R programming software[60]. See dedicated method sections for the details of the snRNA-seq bioinformatic analysis; differentially expressed genes were defined as genes significantly expressed (P adjusted for multiple comparisons < 0.05), and showing, on average, >1.5-fold difference between groups of nuclei in each cell type in every DEG comparison. Volcano plots were constructed by plotting the log2(fold change) of lesion type with smallest p-value for each gene in the x axis and by plotting standard deviation of random (donor) effects for each gene on the y-axis. Statistical analysis used two tailed parametric or non-parametric t-tests for two groups, and Fisher's exact test and one-way analysis of variance with corresponding post hoc tests for multiple group comparisons. Data distributions are presented as barplots, dotplots (with individual data points) and heatmaps. Log CPM gene expression values in the dot plots and heat maps were averaged, mean-

centred, and *z*-score-scaled. Dot size indicates the percentage of nuclei in the cluster in which the gene was detected; among the nuclei in which the gene was detected, the expression level was mean-centred and scaled. Graphical object in Fig. 1a was created with BioRender.com.

Data availability

All raw snRNA-seq data (FASTQ files) will be deposited to EGA. An interactive web browser to analyse cell-type specific expression levels of genes and transcriptomic changes in MS versus control tissue is available at https://malhotralab.shinyapps.io/MS_broad/ (for broad cell types) and at https://malhotralab.shinyapps.io/MS_fine/ (for fine cell types).

Supplementary files

S1: Sample metadata with QC summary.

S2: QC report generated by SampleQC.

S3: Top 100 up- and down-regulated genes for each cell type in GM.

S4: Top 100 up- and down-regulated genes for each cell type in WM.

S5: Table of genes both DE in lesions and significantly associated with GWAS SNPs.

S6: Heatmaps of MOFA factors explaining at least 5% of variance by cell type, for GM and WM, and GSEA results for all factor and cell type combinations in both GM and WM.

Code availability

The source code used to analyse the snRNA-seq data in the current study is available online at https://wmacnair.gitlab.io/MS_lesions_snRNAseq.

Author contributions

DM designed the study and managed the project; WM, CFC, AW, GCB and DM coordinated all the work and wrote the manuscript; WM designed the computational analysis, performed QC, clustering, differential gene expression, cell type composition, PAGA and MOFA+ analysis; DC developed and optimised the snRNAseq protocol; EH assisted with QC and differential gene expression; DC with assistance from SJ performed the nuclei extraction,

library preparation and snRNAseq experiments, and isolated DNA for genotyping, for 173 samples; DC, PK and SJ selected cluster marker genes; SJ and AW characterised WM samples from EBB; SJ performed and analysed the GPR17 validation experiments; SA, MM and EN characterised WM samples from ABB; NS and VO performed MBP stainings and characterization of GM lesions from UKTB with assistance from DC, AW and RM; LF, LC, SS and EU helped with procuring WM and GM tissue samples from ABB and UKTB, and interpretation of the results; JB and EA mapped raw sequencing reads and quantified spliced/unspliced reads; JB quantified ambient RNA, performed genotyping QC, identified sample swaps, performed MAGMA analysis and developed the ShinyApp; EA performed data integration using Seurat and Harmony; WM and MR developed SampleQC method, and MR provided input on the DE analysis; All co-authors read, revised and approved the manuscript.

Acknowledgements

We thank Catharine Fournier Aquino and her team at the Functional Genomics Centre Zurich (FGCZ) for help with illumina sequencing of 10X single nucleus RNA-seq sample libraries; Kelly Bales and Irene Knuesel for their support with initiating this project; Pierre-Luc Germain at the University of Zurich, Switzerland for assistance with scDblFinder analysis; and members of Malhotra lab and Collin lab at Roche for fruitful discussions of the results. Part of the computation/data handling was enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at the Uppsala Multidisciplinary Center for Advanced Computational Science, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. EA has been funded by European Union, Horizon 2020, Marie-Sklodowska Curie Actions, and grant SOLO, number 794689. G.C.-B.'s research group-Swedish Research Council (grant 2019-01360); the European Union (European Research Council Consolidator Grant EPIScOPE, 681893), Ming Wai Lau Centre for Reparative Medicine, Knut and Alice Wallenberg Foundation (grants 2019-0107; 2019-0089), Swedish Society for Medical Research (SSMF, grant JUB2019), Olav Thon Foundation (grant 20-4543), F. Hoffmann-La Roche, and Karolinska Institutet. MDR acknowledges support from the Swiss National Science Foundation (grant No. 310030_175841) and from the University Research Priority Program Evolution in Action at the University of Zurich.

Competing interests

The study was funded by F. Hoffmann-La Roche Ltd. DM, DC, JB, WM, LF, LC, EU and SS are full time employees of F. Hoffmann-La Roche Ltd. The other authors declare no competing interests.

Bibliography

1. Lubetzki, C., Zalc, B., Williams, A., Stadelmann, C. & Stankoff, B. Remyelination in multiple sclerosis: from basic science to clinical translation. *Lancet Neurol.* **19**, 678–688 (2020).
2. Chataway, J. *et al.* Efficacy of three neuroprotective drugs in secondary progressive multiple sclerosis (MS-SMART): a phase 2b, multiarm, double-blind, randomised placebo-controlled trial. *Lancet Neurol.* **19**, 214–225 (2020).
3. Cadavid, D. *et al.* Safety and efficacy of opicinumab in patients with relapsing multiple sclerosis (SYNERGY): a randomised, placebo-controlled, phase 2 trial. *Lancet Neurol.* **18**, 845–856 (2019).
4. Brown, J. W. L. *et al.* Safety and efficacy of bexarotene in patients with relapsing-remitting multiple sclerosis (CCMR One): a randomised, double-blind, placebo-controlled, parallel-group, phase 2a study. *Lancet Neurol.* **20**, 709–720 (2021).
5. Green, A. J. *et al.* Clemastine fumarate as a remyelinating therapy for multiple sclerosis (ReBUILD): a randomised, controlled, double-blind, crossover trial. *Lancet* **390**, 2481–2489 (2017).
6. Jäkel, S. *et al.* Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).
7. Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75–82 (2019).
8. Absinta, M. *et al.* A lymphocyte-microglia-astrocyte axis in chronic active multiple sclerosis. *Nature* (2021) doi:10.1038/s41586-021-03892-7.
9. van der Valk, P. & De Groot, C. J. Staging of multiple sclerosis (MS) lesions: pathology of the time frame of MS. *Neuropathol. Appl. Neurobiol.* **26**, 2–10 (2000).
10. Lassmann, H. Multiple Sclerosis Pathology. *Cold Spring Harb. Perspect. Med.* **8**, (2018).
11. Barkas, N. *et al.* Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nat. Methods* (2019) doi:10.1038/s41592-019-0466-z.
12. Miao, Z. *et al.* Putative cell type discovery from single-cell gene expression data. *Nat. Methods* **17**, 621–628 (2020).
13. Elbaz, B. & Popko, B. Molecular Control of Oligodendrocyte Development. *Trends Neurosci.* **42**, 263–277 (2019).
14. Wolf, F. A. *et al.* PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
15. Bergen, V., Soldatov, R. A., Kharchenko, P. V. & Theis, F. J. RNA velocity-current challenges and future perspectives. *Mol. Syst. Biol.* **17**, e10282 (2021).
16. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
17. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
18. Brooks, M. *et al.* GlmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J.* **9**, 378 (2017).
19. Zoupi, L. *et al.* Selective vulnerability of inhibitory networks in multiple sclerosis. *Acta Neuropathol.* **141**, 415–429 (2021).

20. Meijer, M. *et al.* Epigenomic priming of immune genes implicates oligodendroglia in multiple sclerosis susceptibility. *Neuron* (2022) doi:10.1016/j.neuron.2021.12.034.

21. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).

22. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, (2019).

23. Bryois, J. *et al.* Cell-type specific cis-eQTLs in eight brain cell-types identifies novel risk genes for human brain disorders. *bioRxiv* (2021) doi:10.1101/2021.10.09.21264604.

24. Argelaguet, R. *et al.* MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* **21**, 111 (2020).

25. Köhler, S., Winkler, U. & Hirrlinger, J. Heterogeneity of Astrocytes in Grey and White Matter. *Neurochem. Res.* **46**, 3–14 (2021).

26. Lin, H. & Peddada, S. D. Analysis of compositions of microbiomes with bias correction. *Nat. Commun.* **11**, 3514 (2020).

27. Holley, J. E., Newcombe, J., Whatmore, J. L. & Gutowski, N. J. Increased blood vessel density and endothelial cell proliferation in multiple sclerosis cerebral white matter. *Neurosci. Lett.* **470**, 65–70 (2010).

28. Griffiths, L. *et al.* Substantial subpial cortical demyelination in progressive multiple sclerosis: have we underestimated the extent of cortical pathology? *Neuroimmunol. Neuroinflamm.* (2020) doi:10.20517/2347-8659.2019.21.

29. Chang, A. *et al.* Cortical remyelination: a new target for repair therapies in multiple sclerosis. *Ann. Neurol.* **72**, 918–926 (2012).

30. Tsouki, F. & Williams, A. Multifaceted involvement of microglia in gray matter pathology in multiple sclerosis. *Stem Cells* **39**, 993–1007 (2021).

31. van Horssen, J. *et al.* Clusters of activated microglia in normal-appearing white matter show signs of innate immune activation. *J. Neuroinflammation* **9**, 156 (2012).

32. Melief, J. *et al.* Microglia in normal appearing white matter of multiple sclerosis are alerted but immunosuppressed. *Glia* **61**, 1848–1861 (2013).

33. Huynh, J. L. *et al.* Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. *Nat. Neurosci.* **17**, 121–130 (2014).

34. Schmid, K. T. *et al.* scPower accelerates and optimizes the design of multi-sample single cell transcriptomic studies. *Nat. Commun.* **12**, 6625 (2021).

35. Colombo, E. *et al.* Siponimod (BAF312) Activates Nrf2 While Hampering NFκB in Human Astrocytes, and Protects From Astrocyte-Induced Neurodegeneration. *Front. Immunol.* **11**, 635 (2020).

36. Magliozzi, R. *et al.* Meningeal inflammation changes the balance of TNF signalling in cortical grey matter in multiple sclerosis. *J. Neuroinflammation* **16**, 259 (2019).

37. Gelman, A. Scaling regression inputs by dividing by two standard deviations. *Stat. Med.* **27**, 2865–2873 (2008).

38. Trapp, B. D., Bö, L., Mörk, S. & Chang, A. Pathogenesis of tissue injury in MS lesions. *J. Neuroimmunol.* **98**, 49–56 (1999).

39. Germain, P.-L., Lun, A., Macnair, W. & Robinson, M. D. Doublet identification in single-cell sequencing data using scDblFinder. *F1000Res.* **10**, 979 (2021).

40. Macnair, W. & Robinson, M. D. SampleQC: robust multivariate, multi-celltype, multi-sample quality control for single cell data. *bioRxiv* 2021.08.28.458012 (2021)

doi:10.1101/2021.08.28.458012.

41. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
42. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
43. Fort, A. *et al.* MBV: a method to solve sample mislabeling and detect technical bias in large combined genotype and sequencing assay datasets. *Bioinformatics* **33**, 1895–1897 (2017).
44. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
45. Lun, A. T. L., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res.* **5**, 2122 (2016).
46. Lun, A. T. L. *et al.* EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
47. Bakken, T. E. *et al.* Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS One* **13**, e0209648 (2018).
48. Chen, S. *et al.* Dissecting heterogeneous cell populations across drug and disease conditions with PopAlign. *Proc. Natl. Acad. Sci. U. S. A.* (2020) doi:10.1073/pnas.2005990117.
49. Hesterberg, T. Bootstrap. *Wiley Interdiscip. Rev. Comput. Stat.* **3**, 497–526 (2011).
50. Crowell, H. L. *et al.* muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. *Nat. Commun.* **11**, 6077 (2020).
51. Squair, J. W. *et al.* Confronting false discoveries in single-cell differential expression. *Cold Spring Harbor Laboratory* 2021.03.12.435024 (2021) doi:10.1101/2021.03.12.435024.
52. Korotkevich, G., Sukhov, V. & Sergushichev, A. Fast gene set enrichment analysis. *Cold Spring Harbor Laboratory* 060012 (2019) doi:10.1101/060012.
53. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
54. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer, New York, NY, 2009).
55. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
56. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
57. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
58. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
59. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
60. Team, R. C. & Others. R: A language and environment for statistical computing. (2013).
61. Nakagawa, S. & Schielzeth, H. A general and simple method for obtainingR2from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142 (2013).

62. Wallen, Z. D. Comparison study of differential abundance testing methods using two large Parkinson disease gut microbiome datasets derived from 16S amplicon sequencing. *BMC Bioinformatics* **22**, 265 (2021).

63. Buettner, M., Ostner, J., Mueller, C. L., Theis, F. J. & Schubert, B. scCODA: A Bayesian model for compositional single-cell data analysis. *Cold Spring Harbor Laboratory* 2020.12.14.422688 (2020) doi:10.1101/2020.12.14.422688.

64. Zimmerman, K. D., Espeland, M. A. & Langefeld, C. D. A practical solution to pseudoreplication bias in single-cell studies. *Nat. Commun.* **12**, 738 (2021).

65. Hoffman, G. E. & Roussos, P. Dream: powerful differential expression analysis for repeated measures designs. *Bioinformatics* **37**, 192–201 (2021).