Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

Celia Lerma Martin, M.Sc.

born in: Barcelona, Spain

Oral examination: 6th February, 2025

Mapping cell-specific changes and interactions in subcortical multiple sclerosis lesions through paired transcriptomics

Referees: Prof. Dr. Ana Martin-Villalba

Prof. Dr. med. Lucas Schirmer

Per a la meva familia,
que sempre ha cregut amb mi

Abstract

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system, characterized by complex interactions between immune and glial cells. Despite advancements in understanding MS pathology, the molecular mechanisms that drive lesion formation, progression and resolution remain unknown. In my thesis, I want to explore the disease progression and tissue changes happening in control of subcortical white matter compared to subcortical MS lesions. Specifically, I focused on inflamed chronic active (CA) and non-inflamed chronic inactive (CI) lesions with an emphasis on the lesion rim. My aim is to characterize the molecular and cellular drivers underlying progression of these MS lesions using advanced transcriptomic techniques and properly validating them through several methods, such as single-molecule fluorescence *in situ* hybridization or immunofluorescence.

To achieve this, I opted for a combined approach, creating a paired single-nucleus RNA sequencing (snRNA-seq) and spatial transcriptomics (ST) dataset to generate a high-resolution transcriptomic atlas. This approach provides a high cellular resolution based on gene expression, which allowed the identification of cell communities, as well as the study of cell type-specific drivers of inflammation and cell-cell communication in lesion and non-lesion areas. Among these findings, I discovered a previously unidentified astrocyte subtype localized at the lesion core, characterized by the presence of enlarged, motile-like cilia.

These findings provide new insight into the cellular drivers and mechanisms involved in MS lesion progression and cell-cell communication at the lesion rim, which could help identify potential therapeutic targets for halting the disease. Additionally, the methodological approaches that I have used, offer a valuable accessible framework for other scientists to apply to their own data, potentially uncovering therapeutic targets beyond MS.

Zusammenfassung

Multiple Sklerose (MS) ist eine chronisch entzündliche Erkrankung des zentralen Nervensystems, die durch komplexe Interaktionen zwischen Immunzellen und Gliazellen gekennzeichnet ist. Trotz Fortschritten im Verständnis der MS-Pathologie sind die molekularen Mechanismen, die die Entstehung und Entwicklung von Läsionen steuern, nach wie vor unbekannt. In meiner Dissertation untersuche ich den Krankheitsverlauf und die Gewebeveränderungen in der subkortikalen weißen Substanz mittels Kontrollgewebe und Gewebe mit MS-Läsionen. Dabei vergleiche ich insbesondere die „entzündeten" chronisch aktiven (CA) und die „nicht entzündeten" chronisch inaktiven (CI) Läsionen, wobei der Schwerpunkt auf dem Läsionsrand liegt. Mein Ziel ist es, die molekularen und zellulären Stimuli dieser MS-Läsionen mit Hilfe fortschrittlicher Transkriptomtechniken zu charakterisieren und sie durch verschiedene Methoden, wie Einzelmolekül-Fluoreszenz-in-situ-Hybridisierung oder Immunfluoreszenz, zu validieren.

Um dies zu erreichen, verwende ich einen kombinierten Ansatz und erstelle einen gepaarten Datensatz aus Einzelzellkernsequenzierung und räumlicher Transkriptomik, um einen hochauflösenden Transkriptom-Atlas zu erstellen. Diese höhere Auflösung ermöglicht es mir, Zellpopulationen zu identifizieren, sowie zelltypspezifische Faktoren von Entzündungen und Zell-Zell-Kommunikation innerhalb und außerhalb der Läsionen zu untersuchen. Unter anderem entdeckte ich einen bisher nicht identifizierten Astrozyten-Subtyp, der hauptsächlich im Zentrum der Läsion vorkommt und durch die Bildung von außergewöhnlich langen Zilien charakterisiert sind, welche beweglichen Zilien ähneln.

Diese Ergebnisse bieten neue Einblicke in die zellulären Abläufe und Mechanismen, die am Voranschreiten von MS-Läsionen und an der Zell-Zell-Kommunikation am Läsionsrand beteiligt sind. Dies kann zur Identifizierung potenzieller therapeutischer Ziele für die Behandlung der Krankheit beitragen. Darüber hinaus bieten die von mir verwendeten methodischen Ansätze anderen Wissenschaftlern einen wertvollen Ansatz, den sie auf ihre eigenen Daten anwenden können, um möglicherweise therapeutische Ziele jenseits von MS zu entdecken.

Table of contents

List of figures and tables

Figures

Introduction

Results

Appendix

Tables

Methods

Appendix

Guidelines for reading the thesis

- **Use of *I* and *we*** are thought through in this work. *I* has been used whenever an experiment or an interpretation of it is solely done by myself. *We* has been used when other people participated, for example, in the characterization of the MS lesions (Christian Riedl, Annika Hofmann) or in the creation of the pipeline used (Pau Badia i Mompel).
- **Materials and Methods.** All methods were performed by me unless otherwise specified at the beginning of the section in bold. When a methodology was carried out in collaboration, the collaborator's contributions are clearly indicated with their name at the start of the section, while my contributions are marked with 'I' in bold. Additionally, sections describing methods performed by collaborators are written in passive voice, while sections on my own work use the first person.
- **Solutions mixtures**. Some solutions names might have a superscript letter (e.g LFB^{A}). This indicates that the details of these solution mixtures have been provided as a footnote.
- **Numbering**. The European style has been used for writing numbers (e.g., 1.000 for whole numbers and 6,5 for decimals). However, in the bioinformatics sections, all tool versions and numerical values have been written in the American format (e.g., version 2.0.1 or value of 0.2).

Abbreviations

The following table contains all abbreviations present in this work, ordered in alphabetical order.

Abbreviation	Meaning	Abbreviation	Meaning
AS	Astrocytes	EP	Ependymal
ABC	Avidin-Biotin-Complex	EtOH	Ethanol
BBB	Blood-Brain Barrier	Fig	Figure
BC	B Cells	FISH	Fluorescent ISH
BH	Benjamini-Hochberg	g	Gravitational force
BSCB	Blood-Spinal Cord Barrier	GM	Gray Matter
$CaCl_2$	Calcium Chloride	GM-CSF	Granulocyte-Macrophage Colony Stimulating Factor
CCC	Cell-Cell Communication	GWAS	Genome Wide Association
cDNA	Complementary DNA	h	Hour
CIS	Clinically Isolated Syndrome	H_2O_2	Hydrogen Peroxide
CNS	Central Nervous System	H&E	Hematoxylin and Eosin stain
CSA	Cross-Sectional Area	HCl	Hydrochloric acid
CSF	Cerebrospinal Fluid	HDST	High-Definition Spatial Transcriptomics
CTRL	Control	HRP	Horseradish Peroxidase
DAB	3,3'-Diaminobenzidine	IF	Immunofluorescence staining
DAMP	Damage-Associated Molecular Pattern	IHC	Immunohistochemistry
DBiT-seq	Deterministic Barcoding in Tissue for spatial omics Sequencing	ISH	*In Situ* Hybridization
DEG	Differentially expressed genes	ISS	*In Situ* Sequencing
DEPC	Diethyl Pyrocarbonate	iST	Imaging-Based ST
dH_2O	Distilled Water	ITBS	Intercept Tris Buffer Saline
Dis	Disease	LAS X	Leica Application Suite X
DTT	Dithiothreitol	LC	Lesion Core
EBV	Epstein-Barr Virus	LFB	Luxol Fast Blue
EC	Endothelial Cells	LPS	Lipopolysaccharides
EDTA	Ethylenediaminetetraacetic acid	LR	Lesion Rim
MC	Myeloid cells	PPWM	Peri-Plaque White Matter

MDS	Multidimensional scaling	Prolif	Proliferative
MeOH	Methanol	RCA	Rolling Circle Amplification
MERFISH	Multiplexed error-robust FISH	RCP	Rolling Circle Products
$Mg(Ac)_2$	Magnesium acetate	React	Reactive
MHC	Major Histocompatibility complex	Rev	Revision
min	Minutes	RIN	RNA Integrity Number
MOGAD	Myelin Oligodendrocyte Glycoprotein Antibody-associated Disease	RNA-seq	RNA-Sequencing
MRI	Magnetic Resonance Imaging	ROI	Region Of Interest
mRNA	Messenger RNA	RPM	Revolutions Per Minute
MS	Multiple Sclerosis	RT	Room Temperature
MS-A	MS Acute lesion	SAGE	Serial Analysis of Gene Expression
MS-CA	MS Chronic Active lesion	SC	Stromal Cells
MS-CI	MS Chronic Inactive lesion	scRNA-seq	Single-Cell RNA-Sequenecing
NEU	Neurons	sec	Second
NGS	Next-Generation Sequencing	seqFISH	Sequencing FISH
NMO (SD)	Neuromyelitis Optica (Spectrum Disorder)	smFISH	Single-Molecule RNA FISH
OL	Oligodendrocytes	snRNA-seq	Single-nucleus RNA-Sequencing
ON	Optic Neuritis	sST	Sequencing-based ST
OPC	Oligodendrocyte Progenitor Cell	ST	Spatial Transcriptomics
PAS	Periodic Acid Schiff	Stereo-seq	Spatio-Temporal Enhanced Resolution Omics Sequencing
PBS	Phosphate-Buffered Saline	TBB	Turnbull Blue staining
PBS-T	PBS with Tween	TC	T Cells
PCA	Principal Component Analysis	TransC	Transitioning Ciliated
PCR	Polymerase Chain Reaction	ULM	Univariate Linear Model
PFA	Paraformaldehyde	UMAP	Uniform Manifold Approximation and Projection
Phago	Phagocytic	UMI	Unique Molecular Identifier
PIXEL-seq	Polony-Indexed Library Sequencing	VI	Vascular Infiltrating
PLP	Padlock Probes	WM	White matter

1. Introduction

Multiple sclerosis (MS) is the most prevalent chronic inflammatory disease of the central nervous system (CNS). It is characterized by inflammation and demyelination of the white matter of the brain and spinal cord, which can lead to varying levels of axonal loss and astrogliosis[1]. MS is generally considered an autoimmune disorder, more common in females, marked by the infiltration of peripheral immune cells into the CNS, as well as activation of the innate immune system involving brain-resident microglial cells[2]. A recent study has identified the first prospective autoantigen-specific biomarkers during the preclinical phase of MS, which continue to be present throughout the disease course[3]. This finding is promising, as it could pave the way for the discovery of a disease-specific autoantibody, similar to those found in conditions like neuromyelitis optica spectrum disorder (NMOSD) and myelin oligodendrocyte glycoprotein antibody-associated disease (MOGAD)[4]. However, the classification of MS as an autoimmune disease remains disputed due to the absence of a definitive autoantigen[5]

In the later stages of MS, the disease is characterized by compartmentalized inflammation that contributes to ongoing degenerative changes in the CNS[5–7]. This progressive neurodegeneration can result in a wide array of symptoms, and many affected individuals eventually experience impaired mobility and cognition[8].

Currently, there is no cure for MS, nor is there medication available to prevent or reverse its progressive neurological deterioration. Existing treatments are disease-modifying medications that aim to reduce the frequency of relapses and limit the accumulation of lesions seen on magnetic resonance imaging (MRI)[8].

1.1 Epidemiology

MS is a complex disease with an unknown etiology. Many studies have tried to identify a single mechanism responsible for MS, hoping to develop therapeutic strategies that would be suitable for all patients. However, MS is multifactorial, involving a complex interplay of genetic and environmental factors[5,8–10]. It remains unclear which phenomena are causative and which are associated effects, complicating both clinical and pathological classifications[9].

In order to understand MS, it is important to recognize that the disease typically develops in genetically susceptible individuals who are exposed to additional environmental risk factors. Although the prevalence of MS in the general population is relatively low, around 0.1%, a genetic

predisposition increases the likelihood of developing the disease, especially in individuals with a family history, where the risk rises to 2-4%[8]. Genome wide association studies (GWAS) have identified over 200 gene variants associated with an increased risk to develop MS. These variants primarily affect immune-pathway genes, with the most significant being the HLA DRB1*1501 haplotype. However, there has not been a genetic risk factor validated so far that influences the clinical course of the disease[8].

Environmental factors also play a significant role in MS, but their impact can vary. Vitamin D is an important factor, where higher levels are advantageous both before the disease develops and throughout its progression. Increased sunlight exposure, which boosts vitamin D levels and is considered beneficial due to the immunomodulatory effects of vitamin D on both the innate and adaptive immune systems[11–13]. Conversely, certain risk factors, which can be present years before disease onset, are associated with an increased predisposition to MS. These risk factors include childhood obesity[14], cigarette smoking[15] and infection with Epstein-Barr virus (EBV), which can lead to mononucleosis in the adolescent and young adults[8,16]. It has been suggested that EBV may cause MS in genetically susceptible individuals by infecting autoreactive B cells[17], potentially causing MS or MS-related diseases, although this has not been definitively proven.

1.2 Anatomical distribution of lesions

Initially, MS pathology was believed to be limited to the brain's white matter, where usually demyelinated plaques are observed. However, advances in modern techniques, such as immunohistochemical staining and MRI, have revealed that lesions also occur in the gray matter. In fact, gray and white matter lesions can be found in various parts of the CNS, including the optic nerve, spinal cord, brainstem, cerebellum, and cerebrum (**Fig. 1**).

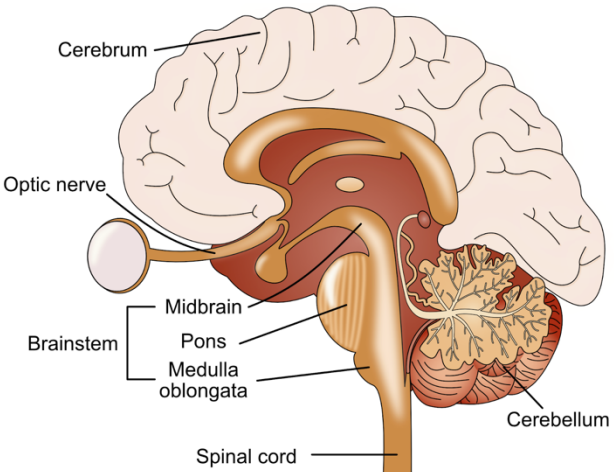


Figure 1. Schematic drawing of the central nervous system core structures.

1.2.1 Optic nerve

The optic nerve is a common target in MS, often presenting without any prior history of demyelinating events[18]. Structural and functional abnormalities in the visual system typically manifest as optic neuritis (ON) in the earliest stages of MS[19]. ON is an acute inflammatory

demyelinating disorder of the optic nerve, characterized by unilateral, subacute and painful visual loss, usually without systemic or other neurological symptoms and predominantly affects young females[18]. Its etiology remains unknown but the emergence and progression of ON into MS involves various pathophysiological mechanisms, including inflammation, demyelination, and axonal degeneration[19,20]. ON can appear as a clinically isolated syndrome (CIS), representing the initial attack of MS in approximately 15-20% of patients[21–23] and occurring in about 50% of patients as part of the disease course[24]. ON can also be associated with other CNS demyelinating disorders like neuromyelitis optica (NMO)[18].

1.2.2 Spinal cord

The spinal cord lies within the spinal column and extends from the brainstem to the lower back[25]. Unlike the brain, the white matter in the spinal cord surrounds the gray matter, which interfaces with the central canal. This arrangement, similar to how the brain white matter interfaces with the ventricular system, suggests a potential link between cerebrospinal fluid (CSF) interfaces and lesion localization[26]. Additionally, the blood-spinal cord barrier (BSCB) is more permeable than the blood-brain barrier (BBB), especially concerning cytokines, which may have implications in various neurological conditions[27]. The increased permeability, along with its anatomical interface, might contribute to the highly inflammatory perivascular and circumferential demyelination seen in MS[28]. This inflammation affects both gray and white matter, accounting for approximately 80% of MS lesions detectable by MRI[29,30], and leading to spinal cord atrophy and neuroaxonal degeneration[31].

Lesions allocated in the spinal cord are more symptomatic than those in the brain, frequently leading to significant impairments such as difficulties with walking, coordination and incontinence, causing significant clinical disability. Importantly, it has become clear that brain and spinal cord lesions correlate poorly, suggesting that the disease progresses independently in these CNS regions[26,32–34].

1.2.3 Brainstem

The brainstem is a critical part of the CNS that connects the cerebrum to the spinal cord and cerebellum. It consists of three main parts: the midbrain, pons and medulla oblongata. MS patients frequently experience lesions in the brainstem along the disease course, with the pons being the most commonly affected area, accounting for approximately 46% of such lesions, followed by the midbrain and medulla oblongata[35].

The pons is responsible for regulating unconscious processes such as breathing and sleep cycles. Due to its significant role and high susceptibility to MS lesions, symptoms arising from the brainstem can be diverse and severe. The brainstem also houses the cranial nerves, which may contribute to the patients experiencing symptoms like difficulties in controlling eye movements, leading to blurred or double vision, facial sensory disturbances and weakness, coordination problems and vertigo[36].

1.2.4 Cerebellum

The cerebellum is a brain region essential for coordinating tasks and fine movements as part of the sensory-motor network, and it also plays a vital role in cognitive-behavioral systems. As a common site for MS lesions[37], damage to the cerebellum in MS patients results in a range of severely debilitating motor and cognitive impairments that significantly affect quality of life and often indicate a poor prognosis[38].

The primary cerebellar symptoms include, but are not restricted to, tremor, nystagmus (repetitive, uncontrolled eye movements), and scanning speech (an abnormal speech where each syllable is pronounced slowly, with pauses between them)[36].

1.2.5 Cerebrum

Within the cerebrum, lesions are categorized into two main types: cortical or gray matter lesions, and subcortical or white matter lesions where the primary emphasis of this work will focus on.

1.2.5.1 Gray matter lesions

Demyelination can occur in both white and gray matter as both contain myelin[39–41], but cortical lesions are less inflammatory that those present in white matter[42]. Additionally, cortical demyelination can occur independently of white matter pathology, suggesting it could be an early or initial target in the progression of MS[43,44]. Approximately half of the cortical lesions are perivascular, however, they exhibit substantially less permeability in the BBB compared to white matter lesions[42].

Subpial cortical lesions account for 60% of cortical lesions[45] and originate from the pial surface of the brain due to meningeal inflammation[44]. These lesions affect all six cortical layers with only marginal involvement of the white matter and are unique to multiple sclerosis, not observed in other inflammatory brain diseases[42,46]. While this is the most widely accepted definition[44], although different descriptions exist[39,47], highlighting the still incomplete understanding of their origin and

evolution. Some meningeal infiltrates resemble lymphoid follicular structures[48–50] and show immunoreactivity for Epstein-Barr virus components[51]. These inflammatory aggregates are believed to contribute to both cortical demyelination, MS disease progression, and increased vascular permeability, though the latter effect appears to be more chronic[44,52].

This type of inflammation is prominent not only in chronic MS but also in early MS. In the early stages, both diffuse meningeal and focal perivascular inflammation resulting from the BBB leakage are significantly associated with cortical lesions[41], and are linked with particularly severe demyelination of the upper cortical layers[53]. It has been speculated that early inflammation in the meninges, together with the production of inflammatory cytokines in the subarachnoid space, may initiate cortical demyelination though subpial lesions and set the stage for subsequent subcortical white matter inflammation and demyelination[41,44].

1.3 White matter lesions

White matter (WM) lesions are the most common type of lesion in MS and are a key diagnostic feature of the disease. Focal WM demyelination is usually paired with the destruction and loss of oligodendrocytes, while axons remain relatively intact[1]. This pathology arises from the complex and dynamic interplay between the immune system and glial cells, including myelinating oligodendrocytes and their precursors, microglia and astrocytes[10]. The resulting lesions show significant variability, suggesting a pathogenic heterogeneity in demyelination and axonal destruction, both among different MS patients and between lesions within the same patient[1].

A crucial step in the development of WM lesions is the disturbance of the BBB. In early MS lesions located in the WM, increased permeability at the lesion sites leads to perivenous demyelination, allowing lymphocyte and plasma cell infiltrates to diffuse and trigger an inflammatory reaction around the postcapillary venules and veins. Over time, these lesions can fuse into confluent demyelinated plaques[1]. This perivenous demyelination, together with oligodendrocyte loss, distinguishes MS from other diseases with focal white and gray matter lesions[1,8]. As the disease progresses, signs of oligodendrocytes dystrophy are observed, reflected by impaired myelin protein expression and changes in their distal processes[10]. At the same time, astrocytes are activated during tissue injury, contributing to the formation of gliotic scars. Astrocytes, however, can play a dual role: they can enhance inflammation by releasing effector molecules, but also limit damage by providing metabolic support to axons and maintaining the BBB[54].

This broader inflammatory process drives demyelination and neurodegeneration through various mechanisms involving reactive oxygen and nitrogen species[55,56], activated complement components[10], cytotoxic and antiinflammatory cytokines[57], as well as chemokines and their receptors[58] that through adhesion molecules facilitate the recruitment of different lymphocyte subsets and monocytes [59–61]. These factors, originating from both the innate and adaptive immune systems, can trigger inflammatory demyelination through a cascade where myeloid cell activation and adaptive immunity reinforce each other[5,10].

1.3.1 Innate immunity

Innate immune cells, particularly blood-borne monocytes, macrophages and tissue-resident microglia (collectively known as myeloid cells), play a crucial role in the pathogenesis of MS. Once activated, these cells are pathologically indistinctive[62].

Microglia, the primary endogenous phagocytes of the CNS, are present from the onset of lesion formation and are essential in driving the chronic progression of MS lesions. They are key to the compartmentalization of lesions, often marked by the formation of the lesion rim. Their activation, often triggered by factors such as fibrin precipitation or other local stimuli, can lead to the production of toxic molecules reported to contribute to oligodendrocyte loss in MS[10]. However, their precise role remains complex, as microglia can mediate inflammation as well as promote tissue repair by clearing myelin debris[63].

Monocytes, on the other hand, are recruited in response to chemokines and cytokines released by glial cells and adaptive immune cells, such as T cells, in MS lesions[64]. Upon activation, they differentiate into macrophages that exhibit an innate scavenger phenotype, enabling the rapid removal of damaged myelin from the tissue[65]. During this process, myeloid cells can shift to a proinflammatory state, acquiring antigen-presenting capacities[5]. This shift promotes the expression of MHC-II and co-stimulatory molecules, along with secretion of proinflammatory cytokines and neurotoxic molecules[66].

These inflammatory actions lead to the reactivation of CNS-infiltrating T cells, resulting in neuroinflammation and demyelination. Furthermore, their proinflammatory state of myeloid cells suppresses Treg cell expansion, thereby inhibiting antiinflammatory processes[67]. Despite this, myeloid cells also contribute to the CNS repair by secreting anti-inflammatory cytokines and neurotrophic factors. This suppresses disease promoting activity of astrocytes and autoreactive

T cells, while phagocytosing myelin debris to promote tissue repair and support efficient remyelination[68,69].

1.3.2 Adaptive immunity

The etiology of tissue damage in MS remains controversial, as it remains debated whether the inflammatory process is the primary driving force or if lesions are initiated by a neurodegenerative process that is modified or amplified by inflammation[70]. Nonetheless, it is widely accepted that adaptive immunity plays a critical role. Both helper (CD4+) and cytotoxic (CD8+) T cells take part, although CD8+ T cells outnumber T helper cells in histopathological studies[65,71]. Additionally, T cells reactive to myelin antigens are found in similar numbers in both individuals with MS and those without the disease. This suggests that the key issue in MS may not be the presence of these autoreactive T cells but rather a dysfunction in the regulatory mechanisms that normally keep these cells in check[72]. In line with this, treatments that limit T cell access to the CNS have shown promise in reducing or eliminating new MS lesions[8], highlighting the importance of hindering these cells from transmigration over the blood-brain barrier into the brain.

For many years, it has been observed that the CSF of most MS patients contains unique antibodies known as oligoclonal bands, which are produced within the CNS[10]. Evidence suggests that the antibody-producing function of B-lineage cells plays a crucial role in MS lesions. Recent studies have identified granulocyte-macrophage colony-stimulating factor (GM-CSF)-producing B cells increased in MS patients and might further enhance the pro-inflammatory response of myeloid cells during MS[73]. B-cell depleting therapies using antibodies like anti-CD20 reduced their number and modulated both their cytokine profile and interaction with other immune cells[73]. These studies and therapies have demonstrated significant success in reduction of clinical relapses and suppressing inflammation underscoring the importance of B cells in MS pathology[5].

1.3.3 Lesion staging

The formation of WM lesions in the MS brain is a dynamic process (**Fig. 2**). To assess lesion staging, myelin destruction and the presence of myeloid cells are the most reliable temporal indicators[9]. As myeloid cells degrade myelin in their lysosomes, the resulting degradation products provide insights into the time interval between active destruction and pathological analysis. Additionally, the gradual removal of myeloid cells from lesions means that their presence, absence, or distribution can indicate the age of a given lesion[9].

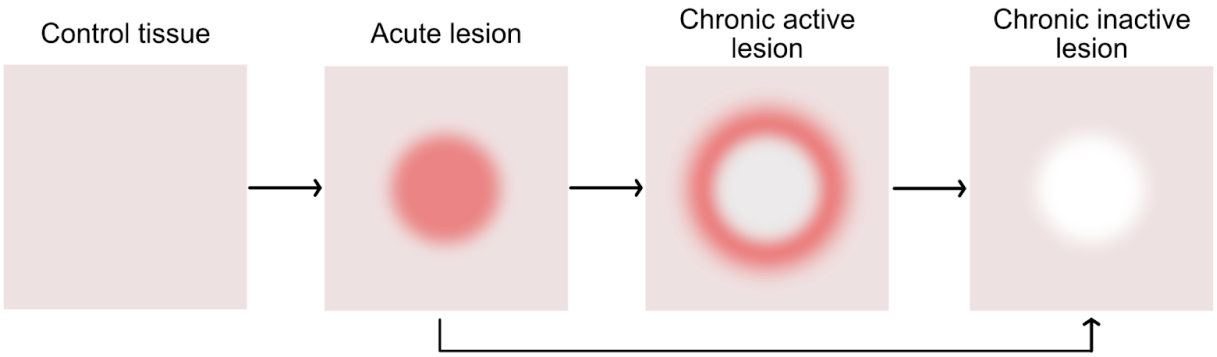


Figure 2. Schematic drawing of possible lesion progression pathways. Red indicates areas of inflammation, while white centers represent the resulting demyelinated cores.

1.3.3.1 Acute Lesions

Acute lesions appear at the earliest stages of WM demyelination and are characterized by focal inflammation and glial reaction. These lesions exhibit hypercellularity, marked by high levels of inflammatory cells and initial myelin breakdown[65]. At the same time, there is a diffuse and dense infiltration of myeloid cells, some presenting a foamy morphology. The presence of foamy macrophages/microglia within a lesion indicates its recent formation, typically days or weeks before the pathological analysis, although it does not confirm an ongoing destruction of the myelin sheath[9]. Additionally, myelin sheaths appear swollen due to intramyelinic edema and begin to undergo partial demyelination, while oligodendrocytes show signs of acute irreversible damage[65]. Plaques forming within these lesions present astrocyte activation, characterized by increased *GFAP* expression, along with microglia activation and occasional caspase-independent oligodendrocyte apoptosis[10].

Several groups have attempted to subclassify this lesion stage based on certain patterns, in order to achieve a more precise characterization[9,10]. However, the controversy regarding whether these patterns represent different etiologies or just reflect temporal stages of MS lesion development still remains. Following this lesion state, the inflammation may resolve becoming inactive or remain active over time, eventually becoming chronic.

1.3.3.2 Chronic Active Lesions

Once lesions become chronic, they develop a distinctive hypocellular and demyelinated core, almost completely depleted of myeloid cells and populated with hypertrophied astrocytes[9,65]. Surrounding this core there is an inflamed rim which varies in thickness, and contains a high density of activated microglia and macrophages featuring myelin degradation products in their cytoplasm, also called foamy macrophages[1,9]. This area appears highly inflamed, given that there

is a high presence of T cells and foamy macrophages, which are associated with active demyelination and tissue injury[1,65]. Outside this area, in the periplaque WM and even reaching into the normal appearing WM, there is some presence of microglia activation partly surrounding degenerating axons[74].

Lesions containing myeloid cells with myelin degradation products in their cytoplasm are often referred to as smoldering or slowly expanding lesions, likely reflecting their ongoing demyelinating activity[74,75]. In some cases, sustained demyelination activity leads to iron accumulation at the rims, which can be visualized through MRI[8], although the source of the iron remains uncertain, with some hypotheses suggesting it might come from phagocytosed oligodendrocytes[76,77]. It remains unclear whether this demyelinating activity contributes to the slow expansion of the lesion, represents a new wave of inflammation and demyelination, or is the last remnant of an earlier demyelinating lesion[9]. Importantly, the presence of a rim of myeloid cells does not necessarily indicate that active myelin destruction is occurring at the time of analysis, but rather that these cells may persist after the initial demyelinating process.

1.3.3.2.1 Remyelination events

Understanding the factors that influence lesion behavior, such as resolution of inflammation, persistence of smoldering lesions, or initiation of remyelination, represents a challenge to the field. Remyelination is the process of forming new myelin sheaths or repairing damaged ones around axons. It can occur at various stages of lesion formation, from early onset to chronic disease stages[78], and is therefore not typically classified as a separate lesion category[9]. Despite its frequency, the factors determining its likelihood to happen are not well understood.

Remyelination events in MS are influenced by a combination of systemic factors such as disease chronicity, sex and age, as well as local variables[78–81]. Lesion location significantly impacts remyelination, with subcortical lesions often exhibiting extensive remyelination compared to periventricular or cerebellar lesions[82,83]. However, its extent may even differ between lesions in the same patient[78]. Locally, myelin repair is affected by the persistence of a baseline level of inflammation, influenced by the composition of the inflammatory infiltrates, which impacts the repair capability within the lesion[84–89]. Additionally, the presence of oligodendrocyte progenitor cells (OPC) is crucial, as they play a central role in remyelination. Often found arrested at the plaque edge[90], OPCs have the potential to differentiate into myelinating oligodendrocytes upon exposure to demyelinating stimuli[91]. However, these myelinating oligodendrocytes may exhibit abnormal phenotypes such an increased size and the absence of their usual perinuclear halo[65]. While they

can initiate normal remyelination, the resulting myelin sheath may be thinner than usual or exhibit aberrant wrapping around the axons[9]. Moreover, dysfunctional oligodendrocytes may contribute to tissue damage by failing to provide trophic support to axons, although their precise role in repair remains uncertain[8,65]. In rare cases of lesion with pronounced astrocyte damage, remyelination may also be facilitated by Schwann cells[92,93].

This process is frequently impaired in most chronic MS lesions, with only about 20% displaying extensive remyelination, often referred to as shadow plaques[83,94]. These shadow plaques can form through the entire lesion or cover only focal areas within the plaque, particularly at the lesion borders[78]. Yet, it remains unclear whether remyelinated areas have an increased susceptibility to new waves of demyelination due to the potential instability of the newly formed myelin[95,96]. Additionally, it is challenging to determine when new myelin sheaths are formed or if ongoing remyelination is occurring[9].

1.3.3.3 Chronic Inactive Lesions

Over time, lesions progress into a chronic inactive stage, which is the most abundant type of lesion in MS[1]. They exhibit a clearly demarcated lesion core that stands out from the surrounding normal-appearing WM[65]. Despite some partial axonal preservation, the core displays a marked loss of axons[97] and scattered axonal swelling, suggesting ongoing axonal damage, potentially of non-inflammatory nature[9].

In chronic inactive lesions, rim-associated inflammation is nearly absent, with a reduced density of macrophages and activated microglia altogether; their levels have been described as even lower than the levels found in the normal-appearing WM[98]. Within the demyelinated hypocellular core, mature oligodendrocytes are nearly depleted and replaced by an abundance of astrocytes that are forming the astroglial scar[1,9,65]. The development of the astrocytic scar within the lesion distorts normal tissue architecture and creates an environment that inhibits axon regeneration and tissue repair[99]. In contrast, at the lesion rim where glial scar formation is incomplete, the tissue architecture is less disturbed[99,100].

1.4 Transcriptomics

The transcriptome was initially defined as '*the complete complement of mRNA molecules generated by a cell or population of cells*' [101]. While hundreds of cell types in the human body essentially share the same DNA, transcriptomics, along with other omics techniques, aims to assess cellular diversity and specific functions at various levels. These techniques include

genomics, proteomics, and metabolomics, each providing complementary insights into cellular processes[102].

The first efforts to profile mammalian transcriptomes began in 1991[103], with methods like serial analysis of gene expression (SAGE)[104] and microarrays[105] dominating the field for a decade[106]. Next-generation sequencing (NGS) then revolutionized the field, surpassing microarray analysis with the first RNA-seq papers published in 2006[107]. Yet, RNA-seq required high-throughput to compete against microarrays, and that was achieved in 2008 with short-read technology[108–110]. This advancement allowed for unique mapping of fragmented cDNA reads to the genome, improving the identification of spliced reads and enabling de novo transcriptome assembly with a reference genome[106]. By 2009, the entire repertoire of mRNA from a single cell, known as the single-cell transcriptome, was described[111].

In contemporary multidisciplinary projects, global transcription profiling is often the first omics technology used, providing insight into gene expression levels and helping to generate hypotheses[112]. This approach is particularly relevant in the study of MS, where transcriptional studies have been conducted to decipher the disease's complex pathology. Previous studies had set out to generate human MS atlases[113,114] in order to provide the community with new insights as well as high-quality data. This enables researchers to conduct their own studies, test hypotheses, and advance our understanding of the disease. However, it often requires trade-offs to overcome its limitations. Therefore, this initial step is followed by additional omics technologies for targeted, hypothesis-driven work, ultimately characterizing specific genes and proteins within tissues of interest[112].

1.4.1 Single-cell RNA-sequencing overview

Recent technological advancements have led to the development of single-cell RNA-seq (scRNA-seq)[115], a technique that has transformed molecular biology. By enabling transcriptome profiling at an unprecedented scale and resolution, scRNA-seq allows for the study of dynamic changes and heterogeneity in eukaryotic cells[116].

Typically, this method requires the release of intact and viable single cells from tissues, to be then individually labeled using microfluidic, droplet-based, or limiting dilution methods in order to profile the mRNA molecules per cell. A notable exception is single-nucleus RNA-seq (snRNA-seq), in which nucleus from tissues are mechanically recovered and processed[117]. Various protocols exist in this field, differing in amplification technology, transcript coverage, and liquid handling

automation in plates[118]. These protocols, along with advances in nanodroplets, picowell technologies, and *in situ* barcoding, have enabled parallel sequencing of tens of thousands of cells[119]. For instance, the Chromium Controller, a widely used device from 10x Genomics, isolates single cells or nucleus in droplets within an oil-based emulsion for mRNA capture, reverse transcription, and molecular and cellular barcoding[120]. Regardless of the method or protocol used, this technology has permitted unbiased, genome-scale assessments of cellular identity, heterogeneity, and dynamic change for thousands to hundreds of thousands of cells[102].

Scaling up technologies to profile large numbers of cells in parallel has been crucial for advancing single-cell transcriptomics. This exponential increase in cell numbers analyzed had to come alongside rapidly evolved computational methods for data processing, quality control, and interpretation[118]. The obtained mRNA sequence reads are mapped to genes and cells of origin using either cellular barcodes or unique molecular identifiers (UMIs) and a reference genome, thereby producing a count matrix of cells by genes. After quality controls, different cellular populations can be identified using literature-based marker genes, and then biologically relevant questions can be investigated[116].

Single-cell transcriptomic was initially applied to model organisms, such as mice, zebrafish, worms etc., but has recently advanced toward creating a comprehensive human body atlas, enabling the assessment of changes in aging, disease, and response to therapeutic treatments[118]. However, scRNA-seq captures only one layer of the regulatory machinery governing cellular function and signaling. To complement it, there have been efforts to measure other modalities at single-cell resolution, such as chromatin accessibility[121], surface proteins[122], T/B cell receptor repertoires[123], and spatial location[124], providing a more comprehensive understanding of cellular biology.

Single-cell omics technologies have revolutionized biomedical research by providing insights into cellular heterogeneity and function[125,126]. However, most high-throughput techniques rely on isolating intact and viable cells from tissue, compromising the spatial context necessary for understanding cellular interactions within structured microenvironments[102]. In response to this challenge, spatial omics emerged, integrating histological imaging and spatial profiling measurements[116].

1.4.2 Spatial transcriptomics overview

Deciphering cellular organization within tissues is crucial, as the spatial arrangement of cells is closely related to their biological functions and provides insights into tissue biology and pathology[127]. Initially, studies characterizing healthy and abnormal spatial organization within tissues often used *in situ* hybridization (ISH) and immunohistochemistry (IHC). These techniques allow the visualization of biological processes by mapping a selected number of DNA, RNA and protein within tissues[127]. Recently, numerous spatially resolved, next-generation sequencing technologies[128–130] for gene transcription in intact tissue, commonly referred to as spatial transcriptomics (ST), have been developed. However, currently available techniques offer a trade-off between the number of genes profiled, which can range from tens to thousands or the whole genome, and the queried tissue area that can vary from a whole tissue domain to a single cell or even finer. As the resolution and sensitivity of spatial transcriptomic technologies improve, the integration with other data modalities will provide an opportunity for better tissue characterization[102].

ST technologies offer an unbiased view of the position of each cell relative to its neighbors and non-cellular structures, providing a valuable resource for observing cell type composition, patterns of gene expression, and the co-localization of different cell states. This high-dimensional dataset can serve as a tissue reference atlas, facilitating hypothesis testing. These findings can then be independently validated using ISH or IHC[127].

The most extensive application of ST technologies to date is to elucidate the cell-type composition of a tissue using datasets typically obtained from a single-cell or single-nucleus RNA-seq reference atlas[112]. These atlases are computationally projected onto the ST dataset, enabling the exploration of tissue architecture and providing information on the spatial distribution of both healthy and diseased tissue[131,132]. By preserving the spatial context of cells, disease-promoting tissue patterning niches can be explored by characterizing their altered transcriptional response to disease[127,133]. In addition, the captured gene expression allows for the identification of cell-cell communication events by defining ligand-receptor pairs among cell types[134], thereby contextualizing some of the cellular function in disease.

Currently, ST methods can be broadly categorized into imaging-based approaches and sequencing-based approaches. However this classification is not always clear-cut, as many methods incorporate elements from both categories[127].

1.4.2.1 Imaging-based spatial transcriptomics

Imaging-based spatial transcriptomics (iST) preserves spatial information within intact cells and tissues by imaging a small number of specific mRNA *in situ* via microscopy[102,127]. iST comprises two main modalities: *in situ* sequencing (ISS) and *in situ* hybridization (ISH). Both of these techniques face a tradeoff between capture efficiency and the number of genes profiled, along with challenges like extended imaging times spanning from hours to days and the generation of massive datasets. These obstacles impact scalability and sample stability across multiple imaging rounds[102].

In the case of ISS, one needs to first design a customized library of padlock probes (PLPs), which are long uniquely barcoded oligonucleotides designed to target specific transcripts. Each PLP has ends complementary to adjacent sequences on the target cDNA, so when the probe hybridizes, its ends are brought together. This configuration allows for direct amplification of the target cDNA by rolling circle amplification (RCA), generating rolling circle products (RCPs) that remain in the tissue and can be visualized through staining and imaging[135,136] (**Fig. 3**). ISS can sometimes be combined with cDNA extraction to enable complementary NGS[137] for deeper transcriptomic analysis, blurring the traditional lines between spatial transcriptomics categories.

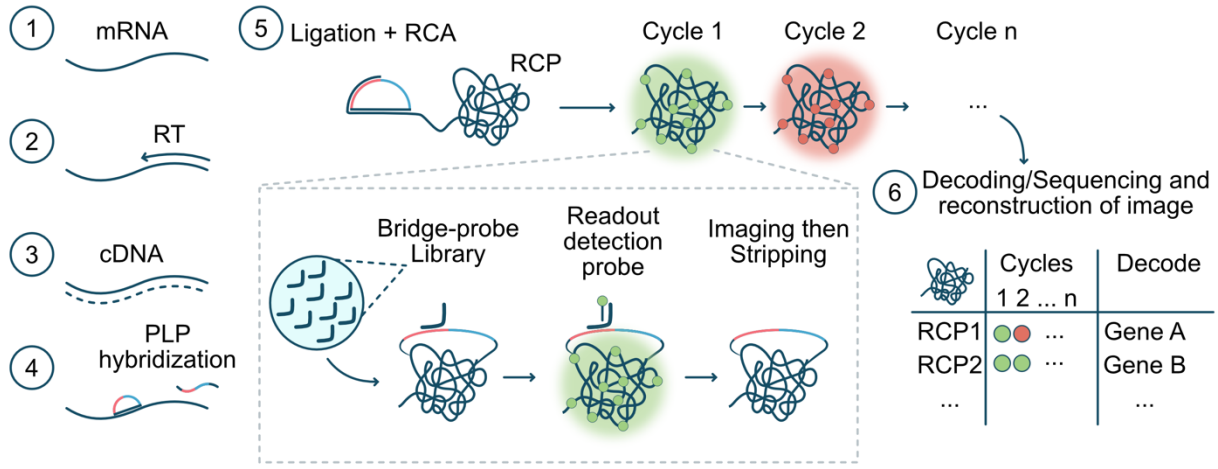


Figure 3. Workflow of *in situ* sequencing (ISS). Adapted from Gyllborg *et. al.* 2020[138]. (1) mRNA is extracted from the sample and (2) reverse-transcribed into (3) cDNA. Next, (4) padlock probes (PLPs) hybridize to their complementary sequences and (5) undergo rolling circle amplification (RCA), generating rolling circle products (RCPs). Different probes then bind to these RCPs and proceed through multiple cycles of staining and imaging. Finally, (6) a decoding algorithm is applied to reconstruct the image based on the acquired information.

On the other hand, ISH involves hybridizing tissue mRNAs or cDNA with complementary labeled, gene-specific probes[102,127]. ISH methods are commonly used, in particular one of its subsets, fluorescent *in situ* hybridization (FISH)[139]. FISH has been used for over two decades, allowing the simultaneous visualization of up to ten distinct fluorescent molecules, depending on the microscope. The most common approach is the single-molecule RNA FISH (smFISH), where each labeled transcript appears as a single spot under microscopy[139,140]. While useful, it requires scalability to increase the number of genes that could be imaged. Recent advancements have enhanced multiplexing capabilities through sequential fluorescence ISH (seqFISH), where unique sets of mRNA are labeled and imaged sequentially, with fluorescent probes being removed after each imaging round[127,141]. Another option is combinatorial multiplexing, such as multiplexed error-robust fluorescent ISH (MERFISH). In MERFISH, successive rounds of hybridizations are imaged to detect the presence or absence of fluorescently labeled probes, and the serial images are decoded using the error-robust barcode associated with each transcript identity[142] (**Fig. 4**).

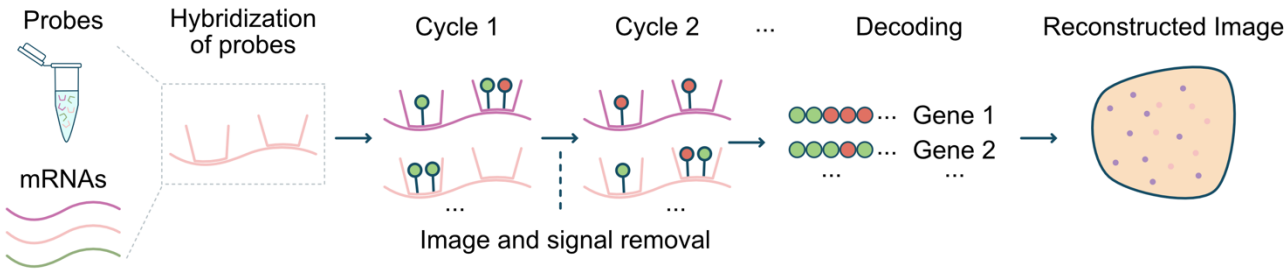


Figure 4. Scheme of MERFISH workflow. Predesigned probes hybridize to complementary mRNA, and unique combinations of fluorescent signals are sequentially imaged and removed in each cycle. These combinations serve as barcodes, enabling the decoding of gene-specific information to reconstruct the image.

Both ISH and ISS methods have advanced significantly in recent years, now capable of detecting around 10,000 genes at subcellular resolution[127]. These improvements have, in turn, driven advancements in three fields: oligonucleotide synthesis, fluorescence microscopy, and single-cell transcriptomics[112]. Firstly, enhanced oligonucleotide synthesis allows for the creation of barcoded hybridization probes[143]. Secondly, the development of sensitive detectors and organic fluorophores enables high-throughput RNA detection[144,145]. Lastly, comprehensive single-cell atlases assist in selecting informative RNA subsets for labeling. Together, these advancements have greatly expanded the capabilities of these methods[112].

1.4.2.2 Sequencing-based spatial transcriptomics

Sequencing-based spatial transcriptomics (sST) consists of extracting mRNAs from a tissue, while preserving spatial information, and then profile it via NGS[102]. The spatial resolution ranges from broad tissue regions to subcellular localization, yet there is a trade-off between sequenced tissue volume and spatial resolution[112], with sST having lower spatial resolution and recovery rates compared to iST[139]. These overall differences in technical performance make each approach suitable for addressing distinct biological questions.

The sST technology relies on an fixed array to profile the entire transcriptome in an untargeted manner for any organism using polyadenylated mRNA[102]. However, transcripts from different cells can be captured at the same spot, requiring a subsequent analysis to identify the cell types present in each spot. Therefore, incorporating an auxiliary tissue stain can verify whether histology and the capture area correspond to the chosen region of interest for analysis[102]. When selecting the area, it is essential to consider variations in RNase content, cell density, extracellular matrix composition, and other features, as they have a direct impact on the capture sensitivity and the amount of RNA captured, two key quality parameters[112].

1.4.2.2.1 Available spatial technologies

Currently, several technologies are available for performing sST, each relying on different methods to generate spatially indexed pixels. One of the most used strategies is array-based mRNA capture, the pioneering method in spatial NGS[128], used in platforms such as Visium[146] (**Fig. 5**).

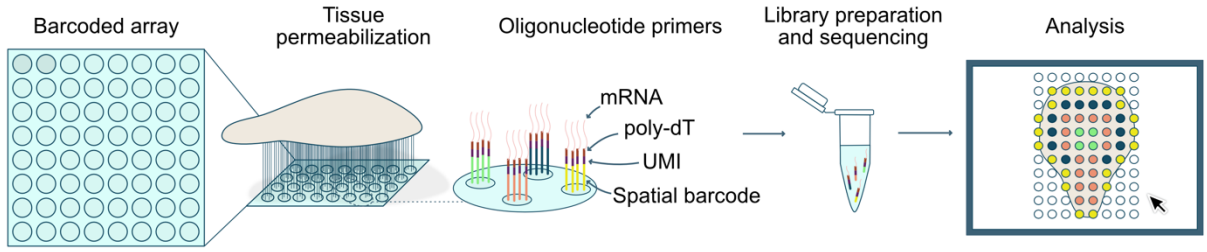


Figure 5. **Overview of array-based spatial transcriptomics.** A tissue section is placed on a barcoded array, and after permeabilization, it releases mRNA that hybridizes to specific oligonucleotide primers. The primers are then detached from the array, libraries are prepared and sequenced, and the data are analyzed to map gene expression across the tissue.

Interestingly, Visium is one of the few techniques that allows for H&E or immunofluorescence auxiliary tissue staining to guide users[102]. In this approach, a microarrayer spotting robot is used to assign a unique barcode to a fixed, predetermined location on the surface of the slide[112]. These 55μm diameter spots, sensitive enough to capture over 10.000 mRNA transcripts, are placed within a 6,5 by 6,5mm capture area and spaced 100μm center to center to prevent liquid mixing during handling[127]. Because of that, Visium recommends an optimization step that aims to determine the ideal duration of tissue permeabilization for mRNA release by converting mRNA to fluorescently labeled cDNA. This is then imaged at different permeabilization times to assess the optimal time that releases the most cDNA with the least lateral diffusion[102].

Another strategy is the one used by the Slide-seq technology[147], which uses solid microparticles for spatial barcoding. Here, 10μm diameter beads, each capable of capturing 500 transcripts, act as the solid support for oligonucleotide synthesis[112]. These beads are randomly barcoded through split-pool cycles and arranged into a tightly packed monolayer on a slide[148]. Their positions are determined by *in situ* indexing prior to tissue mounting, allowing RNA locations to be inferred by sequencing[112]. An improved version, Slide-seqV2, features enhanced barcoding and enzymatic library preparation, recovering approximately 30 to 50% of the transcriptomic information per capture bead compared to droplet-based single-cell transcriptomics from 10x Genomics[102,149]. This allows for the detection of hundreds of thousands of genes per bead. A similar technique, High-definition spatial transcriptomics (HDST)[150], confines beads to tiny wells created on the surface of a slide, rather than on a simple glass slide, achieving a spatial resolution of 2μm.This arrangement ensures that the beads are held in place in a precise pattern, enhancing the spatial resolution and accuracy of RNA location data[102,127].

Moreover, alternative strategies, such as the following methods, are revolutionizing the field with innovative barcode application and enhanced spatial resolution. Deterministic barcoding in tissue for spatial omics sequencing (DBiT-seq)[151] uses microfluidics to apply poly-T barcodes directly to tissue sections, allowing primers to diffuse into the tissue. Spatio-temporal enhanced resolution omics sequencing (Stereo-seq)[152] achieves nanoscale resolution by depositing randomly barcoded DNA nanoballs in an array pattern. Seq-scope[153] offers subcellular resolution by using spatial barcoding to identify the cytoplasm, nucleus, and mitochondria transcripts within cells. Additionally, polony (or DNA cluster)-indexed library sequencing (PIXEL-seq)[154] has significantly increased resolution, achieving up to a 200-fold improvement compared to existing methods.

In all NGS-based methods, after completing surface barcoding and collecting spatially barcoded RNAs, the downstream workflows converge. They proceed with reverse transcription, cDNA amplification, and short-read sequencing to generate reads containing cDNA fragments. These fragments are then mapped to the genome to identify their transcript of origin, resulting in a gene-expression matrix that spatially localizes each detected transcript to its respective pixel[127].

1.4.2.2.2 Omics integration

Studying some genes or rare cell types with sST can be challenging as transcripts from a particular cell may mix with those of neighboring cells rather than being profiled individually[102]. Since many sST techniques lack single-cell resolution, integrating them with scRNA-seq datasets from the same tissue can help address the limitations of both methods. This integration compensates for the lack of spatial information in scRNA-seq and the lack of single-cell resolution in sST[127]. This can be achieved by algorithms that infer the cell type proportion of each spot using single-cell data as a reference[155–157], a process named deconvolution. This process allows for the characterization and interpretation of spots, facilitating the study of spatial features, cell organization, tissue composition, and how the spatial localization of individual cells may impact on tissue microenvironments and their functional dependencies. Additionally, incorporating non-molecular features from histological images can enhance cell identification, molecular resolution, and the detection of spatial patterns of variation[116].

2. Research objectives

In recent years, the MS research field has experienced substantial progress, driven by cutting-edge technologies introduced by the transcriptomic era. These methodologies not only provide enhanced resolution but also generate vast amounts of data, opening up numerous possibilities for analysis. This is particularly true for single-nucleus RNA sequencing, which, having been available for longer, has led the research community to establish widely recognized guidelines and quality standards. There is now a wealth of studies and tools that serve as valuable resources, offering guidance and inspiration for conducting similar analyses. In contrast, spatial transcriptomics is still emerging as a relatively new technology, with fewer established protocols and a lack of consensus on how best to analyze and fully leverage the data it produces. While some analytical methods and applications for spatial transcriptomics are currently being developed, they remain limited in comparison. Therefore, paired atlases that integrate both single-nucleus and spatial transcriptomics are particularly rare, as many studies tend to repurpose existing datasets that may not precisely match their specific disease conditions. This limits the ability to address certain research questions that require the matching of conditions only possible with paired data.

In this context, the aim of my thesis was to leverage a paired transcriptomic atlas to better understand the progression of subcortical WM lesions in humans. By taking advantage of this advanced resolution, I first sought to identify the specific cell types involved in the disease and examine their subtypes to gain insight into how they evolve throughout the progression of the disease.

Additionally, with the spatial transcriptomics data, I aimed to map these cells within the tissue and characterize the formation of disease-specific niches along with their biological properties. This spatial information also would enable the study of cell-cell communication events between neighboring cells within these niches, offering the potential to uncover interactions that could be explored further and shed light on the various processes occurring during disease progression.

3. Materials and methods

3.1 Postmortem human tissue samples

I examined a total of 29 postmortem snap-frozen human brain blocks (21 MS and 8 Control), obtained from autopsies conducted on 16 MS patients and 8 control patients. I categorized samples lacking pathological abnormalities controls, while those from previously diagnosed MS patients underwent further characterization to determine disease stage. All samples were provided by the UK Multiple Sclerosis Tissue Bank at Imperial College London and received ethical approval from the National Research Ethics Committee in the UK (08/MRE09/31). Further details about the donors can be found in Supplementary Table 1.

3.2 Immunohistochemistry

I cut snap-frozen tissue blocks into 16-µm-thick cryosections using a Leica Microsystems CMS3050S cryostat, and then I collected them on SuperFrost plus slides (VWR), which were then stored at -80°C. Immunohistochemical (IHC) techniques used for later histology characterization will be displayed in their own subsection below.

3.2.1 DAB staining

I fixed the sections at RT in -80°C pure MeOH for 5 min, followed by a 5 min wash in 0,1M PBS. Next, I flicked the excess fluid off and the sections and I delineated them with a delimiting Pap pen (Agilent) before blocking them with 10% goat serum (Thermo Fisher) diluted in PBS-T[A] for 1h at RT. Afterwards, I incubated the sections with primary antibodies overnight at 4°C (**Table 1**). The following day, I washed the sections twice for 5 min in 0,1M PBS before I incubated them with biotinylated secondary IgG antibodies for 2h at RT (**Table 2**). After another two 5 min washes in 0,1M PBS, the I incubated the sections for 1h at RT in Avidin-Biotin-Complex (ABC) solution (Vector, 1:500 dilution with PBST), which contains horseradish peroxidase (HRP) conjugated to avidin.

Next, I treated the sections with DAB (3,3'-diaminobenzidine) (DAKO), a chromogen provided as part of a commercial kit, along with a diluent containing hydrogen peroxide (H_2O_2). In the presence of HRP from the ABC complex, DAB is oxidized, producing a brown precipitate that visualizes the antigen-antibody complexes. I applied the DAB solution until I observed a color development

A PBS-T was prepared by mixing 0,1M of PBS and 0,1% of Triton X-100.

under the microscope, with the duration varying between 30 seconds and 1 minute depending on the primary antibody. This way I ensured homogeneous staining comparable across samples. I then stopped the reaction by submerging the sections in 0,1M PBS, followed by counterstaining with 50% hematoxylin for 15 sec. I rinsed the sections rinsed under running tap water, until the running water had no trace of hematoxylin. Finally, I dehydrated the sections. I started with a 5 min incubation in increasing EtOH dilutions (50%→70%→96%→100%) and followed it by two 10 min incubations in 100% Xylene before I mounted them in Eukitt (Orsatec).

3.2.2 DAB-enhanced Turnbull Blue staining

The Turnbull Blue (TBB) staining was performed by the Hametner Lab. Initially, sections were left to dry and then fixed with acetone at 4°C for 10 min. Following this, they were immersed in a 10% ammonium sulfide solution for 1,5h at RT and then dipped 3 to 5 times in dH_2O. Next, sections were incubated for 15 min in a solution comprised of two parts: part one consisting of 20% potassium ferricyanide and part two being 1% HCl (1:1, diluted in dH_2O). Once the incubation was finished, they were dipped additionally 3 to 5 times in dH_2O before being incubated for 1h in a solution containing MeOH (100ml), 0,01M sodium azide (0,065g) and H_2O_2 (1ml). Afterwards, sections were washed in 0,1M Sorensen phosphate buffer and led to a color development reaction with DAB (DAKO) for 20 min. To stop the reaction, sections were immersed in dH_2O and then counterstained for 15 to 30 sec in hematoxylin. Then, they were rinsed under running tap water for 5 min, followed by a differentiation step involving immersion in HCl-EtOH solution[B] . This resulted in sections having a pink color, and by incubating them for 5 min in warm tap water they acquired a blue color, hence termed "bluing". Finally, sections underwent dehydration, starting with a 5-min incubation in an increasing EtOH solution(70%→96%→96%→96%), followed by two 10 min incubations in n-butyl acetate (ester) before being mounted with Epredia Consul-Mount (Thermo Fisher).

3.2.3 Luxol Fast Blue

All the washes I performed in this protocol are carried out by moving the slides 3 to 6 times up and down in the given solution, unless otherwise specified. First, I left sections to dry out for 30 min at RT and then fixed them with 4% PFA for 10 min at RT. After fixation, I washed them twice in dH_2O and were incubated them with 0,1% Luxol Fast Blue (LFB)[C] at 56°C ON. The following

B The HCl-EtOH solution was prepared by mixing 70% EtOH (0,45 mL) with 37% HCl for a final volume of 100mL.
C The LFB solution was prepared by mixing 0,1g of Luxol Fast Blue MSB, 100mL of 95%EtOH and 0,5mL glacial acetic acid.

day, I rehydrated the sections with a series comprising three washing steps in 96% EtOH and three in dH_2O. Next, I immersed them in 0,1% aqueous lithium-carbonate (Morphisto) for 5 min and differentiated them in 70% EtOH until the color reaction developed. This reaction results in an intense blue color in areas with myelin sheaths.

Following this step, I counterstained the sections with the periodic acid-Schiff (PAS) reaction. This consists of incubating the sections for 10 min in 10% Periodic acid, followed by three washes in dH_2O and a 20 min incubation in Schiff reagent (Sigma Aldrich). Next, I washed the sections three times in 2 min washes in a sulfite wash solution[D] and then I rinsed them under running tap water for an additional 10 min. Afterwards, I counterstained the sections with hematoxylin for 1 min and I washed them in dH_2O three times before I differentiated them in HCl-EtOH. Finally, I dehydrated the sections, starting with a 5-min incubation in an increasing EtOH solution(70%→96%→96%→96%), followed by two 10 min incubations in n-butyl acetate (ester) before I mounted them with Eukit (Orsatec).

3.2.4 Immunofluorescence staining

For immunofluorescence (IF) stainings, I started by fixing the sections for 10 min in 4%PFA, ice cold MeOH or acetone depending on the sensitivity of the primary antibody I would use. Next, I flicked the excess fluid off and I delineated the sections with a delimiting Pap pen (Agilent) before I blocked them with 10% goat serum (Thermo Fisher) that was diluted in PBS-T for 30 min at RT. I next diluted the primary antibodies in PBS-T or Intercept-TBS (ITBS) (LI-COR) (**Table 1**), applied them on the sections and incubated them overnight at 4°C. Of note, the antibodies were labeled using cross-adsorbed secondary goat IgG antibodies (H+L) against different host species. The following day, I washed the sections twice for 5 min in 0,1M PBS, and incubated them with secondary antibodies diluted in PBS-T for 2h at RT (**Table 2**). Finally, I washed sections twice for 5 min in 0,1 PBS and mounted them with Fluoromount-G with DAPI.

Table 1. Overview of primary antibodies used for IF and DAB stainings

Target	Species	Fixation	Dilution	Manufacturer	Catalogue	Clone	Staining used
AQP4	Rabbit	Acetone MeOH 4%PFA	1:2500 (PBST)	Sigma Aldrich	HPA014784	-	IF

C The sulfite wash solution was prepared by mixing 20mL of 10% potassium metabisulfite, 500mL of dH_2O and 5mL of 37%HCl.

CD3	Rat	Acetone	1:200 (ITBS)	Bio-Rad	MCA1477	CD3-12	IF
CD11b	Rabbit	Acetone MeOH	1:500 (ITBS)	Abcam	ab133357	EPR1344	IF
CD19	Mouse	Acetone	1:50 (ITBS)	Biolegend	302202	HIB19	IF
CD68	Mouse	Acetone MeOH	1:200 (PBST)	Biolegend	333802	-	DAB
CD163	Mouse	MeOH	1:100 (PBST)	Novocastra	NCL-L-CD163	-	DAB
GFAP	Rat	MeOH 4%PFA	1:200 (PBST)	Thermo Fisher	13-0300	-	IF
MOG	Mouse	MeOH 4%PFA	1:1000 (PBST)	Merck Millipore	MAB5680	8-18C5	DAB
SMI32	Mouse	Acetone MeOH	1:2500 (PBST)	Biolegend	801701	-	IF
SPAG17	Rabbit	Acetone	1:100 (ITBS)	Thermo Fisher	PA5-55912	-	IF

Table 2. Overview of secondary antibodies used for IF and DAB stainings

Target	Host	Dilution	Manufacturer	Catalogue	Staining used
Cross-Adsorbed Biotinylated Ab	Mouse	1:500 (PBST)	Thermo Fisher	62-6540	DAB
Goat anti-'Host' IgG (H+L) Highly Cross-Adsorbed -- Alexa Fluor 488	Mouse	1:500 (PBST)	Thermo Fisher	A-11029	IF
	Rabbit			A-11034	
	Rat			A48262	
Goat anti-'Host' IgG (H+L) Highly Cross-Adsorbed -- Alexa Fluor 555	Mouse	1:500 (PBST)	Thermo Fisher	A-21424	IF
	Rabbit			A32732	
	Rat			A48263	
Goat anti-'Host' IgG (H+L) Highly Cross-Adsorbed -- Alexa Fluor 555	Mouse	1:500 (PBST)	Thermo Fisher	A-21236	IF
	Rabbit			A-21245	
	Rat			A48265	

3.3 Image acquisition

I used a Leica DM6 B brightfield microscope equipped with a K3C camera at a 20x magnification to image DAB, TBB and LFB stainings. I preprocessed the images using Leica Application Suite X (LAS X, v.3.8.1.26810) software and exported them as TIFF files for later post-processing via ImageJ (v.2-2.14.0) software. Additionally, **some brightfield images were also acquired by the Hametner Lab** in Vienna using a Hamamatsu NanoZoomer 2.0HT at a 40x magnification and exported as NPD files. I performed the post-processing of these images using NDP.view2 from Hamamatsu NanoZoomer 2.0HT.

I acquired the fluorescent images using a Leica DM6 B fluorescent microscope equipped with a K5 camera. I preprocessed the images using the LAS X software with the thunder imaging system. I set the focus points at 20x magnification for overview purposes and at 40x magnification for closeups. I performed Z- stacks that consisted in 10 to 20 layers with a step size of 0,7µm across all used channels.

3.4 Lesion type characterization

For the characterization of subcortical white matter lesions, I used a combination of IHC techniques previously described in Section 3.2. I began lesion identification with visualizing myelin loss using LFB) and anti-MOG stainings, as well as assessing myeloid cell activation with anti-CD68 and anti-CD163 stainings. Acute and chronic active lesions both contain high numbers of activated myeloid cells, but they differ in that acute lesions have not yet experienced significant myelin loss due to the recent onset of damage.

In chronic lesions, myelin loss becomes a key feature. I Distinguished lesion types by evaluating myeloid cell activation and detecting iron deposition at the lesion rims, identified by TBB staining. Chronic active lesions present as hypocellular, demyelinated core with distinctly inflamed rims densely populated by myeloid cells and iron deposits. In contrast, chronic inactive lesions exhibit a fully demyelinated core with minimal myeloid cell presence at the rim. **Characterization was supervised by a trained neuropathologist, Simon Hametner, specialized in MS pathology**.

3.5 Sample selection for transcriptomics

I used the RNA Integration Number (RIN) as a criteria for sample selection, with only samples with a RIN value of ≥5,9 being included for both transcriptomic analysis. I cut tissue sections to a thickness of 50µm using a Leica Microsystems CM3050S cryostat, until I collected a final amount

of 10 mg, from which I would extract RNA. I performed the extraction process using the protocols provided by the manufacturer of the RNeasy Mini Kit (QIAGEN), where also TRIzol (Thermo Fisher) and chloroform (Sigma Aldrich) were necessary.

I started by homogenizing the samples using a 22G cannula with 1mL of TRIzol and incubated them for 5 min at RT. Next, I added 200µL of chloroform and vortexed the samples, before I incubated them for another 2 min at RT. I then centrifuged the samples at 10.000g for 20 min at 4°C. After centrifugation, the samples separated into two distinct phases: a clear aqueous phase on top and a denser and opaque phase below. I pipetted approximately 300µL of the aqueous phase into an RNase-free Eppendorf tube and added an equal volume of 300µL of 70% RNase-free EtOH, mixing thoroughly by pipetting up and down. I transferred the contents into a filtered QIAGEN spin tube and centrifuged them at 8.000g for 30 sec, discarding the fluid. Next, I performed a series of cleanup steps as stated by the QIAGEN protocol. For the final step, I added 30µL of RNase-free water directly on top of the filter membrane, followed by centrifugation at 8,000g for 1 min. I repeated this step to ensure optimal purity.

For the last part, I measured RNA integrity using an Agilent 2100 Bioanalyzer with the High Sensitivity RNA assay (Agilent), in accordance with the guidelines provided by the manufacturer. I stored the remaining RNA at -80°C.

3.6 Fluorescence multiplex *in situ* RNA hybridization

For single-molecule fluorescence *in situ* hybridization (smFISH) validation, I used samples that had a RIN value of ≥5,9. I first sectioned the samples from the selected tissue blocks into 16-µm-thick cryosections using a Leica Microsystems CMS3050S cryostat, and then I collected them on SuperFrost plus slides (VWR), which were then stored at -80°C.

I fixed the sections with 4%PFA in two incubation steps: the first for 15 min at 4°C and then the second for 2h at RT. After fixation, I washed the sections twice for 2 min in 0,1M PBS and incubated them for 10 min with H_2O_2 at RT. Next, I washed them again twice for 2 min in dH_2O and boiled them at 200°C in Target Retrieval solution (ACD) for 5 min. Following this, I immersed the sections twice in dH_2O and dehydrated them in 100% EtOH for 3 min before I air dried them. I delineated the tissue areas that I wanted to study later using a hydrophobic ImmEdge pen. Next, I covered the sections with protease (Protease IV, ACD) and incubated them for 30 min at RT. After the protease treatment, I immersed them twice in 0,1M PBS and then incubated them for 2h at 40°C with the mixed target probes (C1:C2:C3; 50:1:1) (**Table 3**) in an RNAscope

hybridization oven (ACD). For the signal amplification and probe channel detection I followed the guidelines provided by the manufacturer. The fluorophores I used for probe labeling were TSA plus Fluorescein, Cyanine3 and Cyanine 5 (Akoya biosciences) with dilutions varying from 1:300 to 1:750. Finally, I mounted the sections using ProLong Gold antifade (Thermo Fisher) and imaged them as outlined in the fluorescent imaging section 3.3.

Table 3. Overview of probes used for *in situ* RNA hybridization

Probe smFISH	Catalogue	Channel	Probe smFISH	Catalogue	Channel
Hs-ADCY2	1241091	C3	Hs-ITGB1	516931	C2
Hs-CD14	418801	C1	Hs-TLR2	403111	C2
Hs-CD163	417061	C2	Hs-SPAG17	834721	C1
Hs-HMGB1	434631	C1	Hs-VWF	560461	C3

3.7 snRNA-seq experimental workflow

3.7.1 Sample preparation

A day prior to the nucleus isolation, I cut the tissue sections to a thickness of 80µm using a Leica Microsystems CM3050S cryostat, collecting a final tissue quantity of 20mg, which I then stored at -80°C.

3.7.2 Materials and buffer preparation

All materials I used for this protocol were RNA free; therefore, before starting the protocol I cleaned with RNaseZap (Thermo Fisher) all benches and utensils that I would use. In addition, I autoclaved the douncers for tissue homogenization one day before nucleus isolation.

The first step of the protocol consists of the preparation of two buffers, which need to be cooled to approximately 4°C and then kept on ice until they are used. The first buffer is the lysis buffer, which I prepared by mixing 5,47g of Sucrose, 250µL of 1M $CaCl_2$, 150µL of 1M $Mg(Ac)_2$, 10µL of 0,5M EDTA, 500µL of 1M Tris-HCl, 17µL of 3M DTT, 50µL of 100% Triton X100, then adjusting the to 50mL with DEPC-treated water.

The second buffer is the sucrose solution, which I prepared by combining 30,78g of sucrose, 150µL of 1M $Mg(Ac)_2$, 17µL of 3M DTT, 500µL of 1M Tris-HCl, then adjusting the volume to 50mL with DEPC-treated water.

3.7.3 Nucleus isolation

Until specified, I conducted all next steps on ice. I started by pipetting 1mL of lysis buffer into the eppendorf tube containing the tissue to ease the initial breakdown. After a couple of times pipetting up and down, I transferred the solution to a pre-chilled glass dounce. I repeated this process for a second time, resulting in a final volume of 2mL. Following this, I homogenized the samples using the douncers, with gentle up-and-down movements to minimize air bubbles and avoid twisting the douncer, which could damage the cells.I limited the strokes to a maximum of 10 per sample before transferring the homogenate to a pre-chilled ultracentrifuge tube, carefully pipetting along the wall of the tube. Next, I slowly and meticulously pipetted 3,7mL of sucrose solution onto the bottom of the tube containing the homogenized tissue, making sure that no bubbles were introduced.

In the next step, I calibrated all the samples to ensure they had the same weight before loading them into the centrifuge. I added lysis buffer as needed to achieve this uniformity. I then loaded them on a SW28 rotor equipped with a swing bucket that could fit up to 6 samples, and centrifuged them for 2,5h at 24.400 RPM at 4°C.

Following centrifugation, I transferred the samples to ice and removed the supernatant via vacuum aspiration using a glass pasteur pipette. Then I carefully popetted 200µL of 0,1M DEPC water-based PBS (Thermo Fisher) onto the bottom of the tube and incubated them on ice for 20 min. Afterwards, I gently resuspended the samples, filtered them twice through 30µm filters (Miltenyi Biotec), and I manually counted the nucleus suspension. My goal was to achieve a final concentration of 1.000 nucleus/µL, so I diluted the samples with 0,1M DEPC water-based PBS to reach this concentration.

3.7.4 Nucleus emulsion and library preparation

Immediately after isolating the nuclei, I loaded them into the 10x Genomics Chromium controller, following the provided guidelines to achieve a recovery rate of 8,000 nuclei per sample. Next, I prepared the libraries following the Chromium Next GEM Single Cell 3' (CG000204 Rev D) protocol from 10x.

During the protocol, I conducted two quality controls to ensure the libraries met the sequencing requirements using TapeStation 4200 analyzer (Agilent) with High Sensitivity D5000 and D1000 ScreenTapes (Agilent) respectively. With the first one, I analyzed the representative cDNA traces to determine the appropriate unique sample index PCR cycles, and for the second one I evaluated

the post-library construction and sample traces to prevent sequencing impurities. If this happened, I performed an additional clean-up step. I sent the finalized libraries to a sequencing facility in Kiel, where they loaded the libraries at a concentration of 300 pM and sequenced them using the NovaSeq 6000 system (Illumina) at a depth of 250 to 300 million reads per sample.

3.8 snRNA-seq computational analyses

3.8.1 Expression count matrix generation

Using the FASTQ files I obtained from sequencing, I generated the expression count matrices using Cell Ranger Count software (v.6.0.2), aligning the data to the GRCh38-2020-A reference transcriptome. I then corrected these count matrices for ambient mRNA using CellBender (v.0.2.0), which I set with the following parameters: model = full; expected_cells = 8.000; total_droplets_included = 50.000; fpr = 0,01; epochs = 150; posterior_batch_size = 5; cells_posterior_reg_calc = 50.

3.8.2 snRNA-seq data processing

For the downstream analyses and data processing for the snRNA-seq datasets, I used Python (v.3.10.0) and the Scanpy toolkit (v.1.9.6) with scripts available at https://github.com/saezlab/VisiumMS and custom made scripts to interpret the data.

3.8.3 snRNA-seq data quality control

I performed independent filtration for each sample to address possible batch effects. I set three parameters for single nucleus filtration steps: one based on genes (> 200 genes), another for mitochondrial gene presence (mitochondrial gene percentage < 5%) and for gene counts (counts < 99th percentile of counts). I kept the genes that were expressed across multiple nucleus (expression in > 3 nucleus) for analyses. Then, I filtered the samples based on the doublet score that I computed with scrublet (doublet score < 0,1). Finally, I normalized raw expression values for each nucleus using the median of total counts (target_sum = None) and log-transformed (log1p).

3.8.4 Data integration and cell annotation

I generated an AnnData object by concatenating (join="outer") all preprocessed nucleus from all samples. Next, I conducted feature selection by computing highly variable genes per sample, followed by the selection of the top 4.096 genes that were identified as variable across the

maximum number of samples. Then, I scaled the genes across the nucleus, and performed a PCA on the selected features. I then used Harmony-py (v.0.09) to integrate the resulting PCs, removing batch effects across samples. Following that, I generated the Nearest neighbors for nucleus by estimating similarities in the PC space (n_pcs = 50). I used the resultant connections to construct a UMAP. Lastly, I used the Leiden graph-clustering method (resolution = 0.25) to cluster the nucleus and then manually annotated them using brain and immune gene markers.

3.8.5 Comparison with other datasets

The first dataset I used in order to validate cell type annotations. For that I compared the generated atlas with another reference human snRNA-seq atlas at the molecular level[114]. I obtained the count matrix and annotation metadata from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180759.

I constructed the pseudo-bulk transcriptomic profiles for both atlases at the cell type level using decoupler-py (v.1.5.1). I filtered the genes based on the following hyperparameters: min_count = 10, min_total_count = 15, min_prop = 0.2, min_smpls = 2. To ensure comparability, I filtered the profiles by identifying the intersection of genes between the two atlases and then performed log normalization using Scanpy (target_sum = None). Finally, I computed the Pearson correlation between the different profiles and adjusted the p-values using the Benjamini-Hochberg correction (BH adjusted $P < 0,05$, $r > 0,75$).

The second dataset I used was a curated list of 125 genes from a recent publication[158], that I used to detect senescent signals. This list is available in Supplementary Data 1 of the publication. Using this gene set, I performed an enrichment analysis on both my atlas and the reference atlas[114], applying the ULM method from decoupler (adjusted $P < 0.05$).

3.8.6 Identification and characterization of cell subtypes

I isolated the major cell lineages from the main snRNA-seq atlas to identify specific cell types within them. I retained the genes that were expressed across enough cell types (number of expressed nucleus > 3), while I excluded samples with not enough cells for a specific cell type (number of cells ≥ 5). I next conducted a PCA on the scaled log-transformed expression values of the most variable genes across as many samples as possible. The number of variable genes I used varied depending on the available cell count: 4.096 genes (number of cells > $1e^4$), 2.048 genes (number of cells > $1e^3$) and 1.024 genes (number of cells < $1e^3$). I did the integration of the PC using Harmony-py, and computed the nearest neighbors per nucleus by identifying similarities

in the corrected PC space (n_pcs = 50). I employed the resulting connections to construct a UMAP manifold and perform clustering with the Leiden algorithm (resolution = 1.0).

I characterized the clusters by using the rank_genes_groups function from Scanpy with the log-transformed counts (adjusted P < 0,05; absolute (log2FC) > 0,5). This function helped identify suitable marker genes that later were used to manually annotate the clusters. The ones lacking clear molecular profiles were excluded (designated as “NA”).

3.9 Spatial transcriptomics experimental workflow

I used the 10x Genomics Visium Spatial Gene Expression platform for all spatial transcriptomics experiments.

3.9.1 Tissue scoring

After carefully assessing histology and RIN integrity measurements, I selected the final regions of interest (ROIs) for the study. Before sectioning the tissue, I designed a custom 3D-printed rectangular prism with sharp edges to delineate the ROIs on the fresh frozen blocks. This ensured that when I made cuts with the cryostat, the tissue sections would already reflect the defined areas of the ROIs. The squared base of the prism measured 6,5 x 6,5mm, to ensure that the ROI would fit inside the capture area of the slides and enabling consistent sectioning of the exact same area for snRNA-seq analysis.

Upon scoring the tissue, I sectioned it into 10µm slices using a Leica CM3050 S cryostat and transferred onto Spatial Gene Expression Slides (PN-1000185) that had been pre-cooled to -22°C inside the cryostat. After I had placed all the tissues on the slide, without foldings, incorrect placements or tissue tearing, I stored the slide in a sealed container at -80°C until further processing. This storage method allowed preservation for up to four weeks while maintaining the sample integrity. No spatial slide reset was necessary (CG000332 Rev B).

3.9.2 Tissue fixation, staining and imaging

The first step of the protocol consisted in fixing the sections. I incubated the slide for 1 minute at 37°C in a PCR machine with a special adapter provided by 10x Genomics, allowing me to perform the incubation without closing the lid. Once the time was up, I immersed the slide in a container of -20°C MeOH for 30 minutes and then continued with H&E staining according to the 10x protocol (CG000160 Rev B). No coverslipping of the sample was necessary before imaging.

Next, I imaged the slides using a 10x magnification lens attached to a Leica DMi8 microscope and pre-processed them using the Leica Application Suite X (LAS X, v.4.13) for general morphological analysis and future spatial alignment. Of note, I placed focus on the fiducial frame rather than the tissue itself. I optimized this imaging process as suggested per imaging guidelines (CG000241 Rev D).

3.9.3 Tissue optimization

This optimization process is essential for quantifying the total mRNA in intact tissue sections before starting the library preparation. It is a one-time requirement for each tissue type of interest. After I scored and sectioned the tissue, I placed it in the capture areas of the Visium Spatial Tissue Optimization Slide. Notably these slides differ from the Gene Expression Slides as they feature 8 capture areas instead of 4, facilitating the assessment of diverse permeabilization times. Once I filled all capture areas, except for one that will be the positive control, I fixed and stained the slide as I previously described, for later permeabilize it for varying times as described in the protocol (CG000238 Rev D).

During this step, mRNA is released and gets reversed transcribed to cDNA that will be fluorescently labeled. Following enzymatic removal of the tissue from the slide, fluorescently labeled cDNA remains, allowing assessment of the optimal permeabilization time via fluorescent microscopy. Based on protocol guidelines, I assessed that the best permeabilization time for subcortical white matter was 18 min.

3.9.4 Slide processing and library preparation

For the library preparation, I placed the tissue samples on the Gene Expression Slides. Following the completion of scoring, sectioning, fixation, staining, imaging and correct assessment of permeabilization time, I generated the libraries according to the established protocols from 10x genomics (CG000239 Rev D). Two essential intermediate steps where I assessed the data quality measurements were done in between this protocol. In the first one I assessed the correct amplification of cDNA using QuantStudio 3 Real-Time PCR (Thermo Fisher) and QuantStudio Design Analysis software(v.1.3). For the second one I determined the appropriate unique sample index PCR cycles by analyzing representative full-length cDNA traces using TapeStation 4200 analyzer (Agilent) with High Sensitivity D5000 ScreenTapes (Agilent).

After library construction, I performed a final quality control to ensure that the samples compiled with sequencing requirements by analyzing their traces to prevent sequencing impurities. For that

I used the TapeStation 4200 analyzer (Agilent) with High Sensitivity D1000 ScreenTapes (Agilent). If this was to happen, I performed an additional clean-up step. Lastly, I sent the libraries to the sequencing facilities in Kiel. Then, they loaded the libraries at a concentration of 300 pM and sequenced on a NovaSeq 6000 system (Illumina) with a sequencing depth of 250 million reads per sample.

3.10 Spatial transcriptomics computational analyses

3.10.1 Demultiplexing and spatial gene count matrix generation

I performed the data demultiplexing using Space Ranger software (v.2.0.0). First, I generated FASTQ files with spaceranger mkfastq, then used these files as input for the spaceranger count pipeline, aligning them with the human reference genome GRCh38-2020-A. Inside the spaceranger count pipeline I was able to integrate the information from previously acquired images, allowing the final spatial gene count matrix to contain data on tissue position per slide. Additionally, this last step generated a Loupe file, compatible with the Loupe Browser (v.7.0.0) from 10X Genomics, enabling initial visualization and exploration of the data without requiring prior quality control steps.

3.10.2 Spatial transcriptomics data processing

For the downstream analyses and data processing for the snRNA-seq datasets, I used Python (v.3.10.0) and the Scanpy toolkit (v.1.9.6) with scripts available at https://github.com/saezlab/VisiumMS and custom made scripts to interpret the data.

3.10.3 Spatial transcriptomics data quality control

I conducted quality control analyses for each slide using Scanpy, filtering genes (< 200 genes per slide) and retaining genes expressed across multiple spots (≥ 3 expressed spots). Next, I normalized raw expression values for each spot by the median of total counts (target_sum = None) and log-transformed.

3.10.4 Cell type deconvolution

I used the cell2location (v.0.1.3) package to compute cell type abundances for each spot. Leveraging the annotated snRNA-seq atlas, I inferred the reference expression signatures of major cell types through regularized negative binomial regressions. I used Hierarchical Bayesian models with hyperparameters N_cells_per_location=5 and detection_alpha=20 to deconvolute

each slide. Post-deconvolution, I computed cell type proportions per spot by dividing the abundance of a specific cell type by the total sum of spot abundances. To assess the efficacy of the deconvolution process, I estimated cell type proportions from each slide and compared them to the actual proportions in corresponding snRNA-seq data using Pearson correlation, both at the sample and cell type levels. Additionally, I determined spatial activity scores using the enrichment method ULM from decoupler with REACTOME pathway gene sets (bandwidth=150).

3.10.5 Manual annotations of distinct tissue areas

Based on the output Loupe file from SpaceRanger software, **myself and Christian Riedl under the supervision of Simon Hametner**, a trained neuropathologist, annotated different areas of the tissue that would later be used for the study. The areas were divided in lesion core, that covered the demyelinated area and had a high amount of astrocyte genes (*GFAP, AQP4*); the lesion rim, that created a distinct border between the demyelinated lesion center and the surrounding myelinated white matter matter, and had a high myeloid cell presence (*CD68, CD163*); the periplaque white matter that had still a high presence of myelin (*MOG, MBP, PLP1*). Each slide was annotated individually based on the same parameters and using Loupe Browser (v.6.3.0) from 10x Genomics.

3.10.6 Generation, characterization and validation of niches

Niche annotation was done in an unsupervised manner across spots for each slide independently. **This first part was performed by Pau Badia i Mompel**, and involved the integration of gene expression and cell type compositions into a multi-view factor model using MOFA+ (v.0.7.0). The transformation of cell type proportions to centered log-ratios was performed using the clr function from the composition-stats python package (v.2.0.0), while gene expression data was summarized into 50 principal components via PCA on the scaled top 4.096 variable genes. These distinct data matrices were used as individual views within the MOFA+ model, from which 15 latent factors were derived. This process used the following specifications: scale_views = False, center_groups = False, spikeslab_weights = False, ard_weights = False, ard_factors = False. Spot-wise nearest neighbors were determined based on similarities computed in the latent factor space. Then, these connections were leveraged to construct a UMAP manifold and perform spot clustering using the Leiden algorithm (resolution=1.0).

Next, I performed manual annotation of clusters into niches based on the presence of specific cell types and pathway activities derived from hallmark gene sets, using the ULM enrichment

method from decoupler. **I identified niche-specific marker genes**, by concatenating the slides into a single AnnData object and the function of Scanpy rank_genes_groups (method = t-test_oversampled_var). Besides characterizing the main cell types per niche, **I calculated the average cell type proportions per slide** and statistically compared them against the rest using the Wilcoxon rank-sum test (adjusted P < 0,05).

Pau Badia i Mompel evaluated the level of overlap between computationally annotated niches and those annotated by a neuropathologist, using the Jaccard index and the adjusted Rand index for each slide, disregarding categories not shared between annotations. Furthermore, similarities between intra- and interniche spots were evaluated through Pearson correlation applied to pseudo-bulked gene expression and mean clr-transformed cell type proportions.

3.10.7 Spatial trajectory analysis across niches

I concatenated the slides and generated pseudo-bulk profiles for each combination of niche and slide using decoupler. I included genes based on the following hyperparameters: group = None, min_count > 10, min_total_count > 15. With the resulting gene profiles, I performed log transformation and normalization (target_sum = $1e^{4}$). Next, I computed the spearman correlation for each gene, following the niche order: WM, PPWM, LR, LC, VI. Then, I used these gene correlation statistics to infer pathway activities that exhibit differential changes along the trajectory, leveraging the PROGENy resource via the ULM method.

Also, I determined significant pathways (P < 0,05) for the pseudo-bulked gene expression profiles. Lastly, I assessed the correlated genes, categorized as positive and negative, or enrichment of gene sets from REACTOME using the ORA method from decoupler. I excluded the gene sets containing terms such as FETAL, INFECTION, or SARS before enrichment computation.

3.11 Characterization of ciliated astrocytes

I excluded the samples and slides containing ependyma, MS549H and MS549T, from the analysis due to the similarity of ependymal cells to ciliated AS. I generated a gene set specific to ciliated AS by filtering marker genes obtained from the filtered atlas using Scanpy (method = t-test_oversampled_var, adjusted P < $1e^{-4}$, log2FC > 1).

I performed an analysis that predicted that spots with a proportion of AS greater than 0,25 contained ciliated AS by computing enrichment scores using the derived gene set with the ULM method from Decoupler (weight = None, P < 0,05, score > 0).

To account for variations in spot counts per slide**,** I determined the niches where cilia AS were found by performing a bootstrap strategy**.** The total number of spots predicted to contain ciliated AS was estimated for each niche, followed by 1.000 random permutations to assess if the sampled ciliated AS counts exceeded the original estimate, yielding empirical p-values per niche (adjusted $P < 0,05$).

I then performed an enrichment analysis, using the ciliated AS gene set against the REACTOME collection using the ORA method from decoupler (adjusted $P < 0,05$). I excluded the gene sets containing terms such as FETAL, INFECTION, or SARS prior to enrichment computation. I assessed the transcription factor activity at both the nucleus and spot levels using the CollecTRI resource with the ULM method from decoupler.

3.12 Multidimensional scaling of samples

The following analyses were designed by Pau Badia i Mompel and Ricardo Ramirez, while the biological interpretation of the resulting data was performed by myself.

Multidimensional scaling (MDS) from sklearn (v.1.4.0) python package was used to capture and visualize sample differences across various dimensions, including cell type proportions, cell subtype proportions, deconvoluted cell type proportions, cell type gene expression, and tissue niche gene expression. To facilitate analysis, cell type, cell subtype, and deconvoluted cell type proportions underwent logarithmic transformation using the clr function from the composition-stats Python package.

For gene expression analysis, latent factors capturing shared gene programs between cell types or niches were inferred using the MOFAcellulaR package for snRNA-seq and ST samples, respectively.

In the case of snRNA-seq data, four multicellular factors were inferred using multicellular factor analysis with MOFA+ on the pseudobulk expression profiles of each cell type and sample. Pseudobulk profiles were constructed by aggregating gene counts from all cells within a given cell type and sample, with profiles containing fewer than 10 cells excluded from the analysis. Moreover, genes expressed in at least 25% of the samples within each cell type were retained, with a gene considered expressed if it had at least 100 counts. Samples within a cell type exhibiting a gene coverage of less than 90% or containing fewer than 10 samples or more than 50 genes were also excluded from analysis.

For ST data, three factors were inferred using multicellular factor analysis on the pseudobulk expression profiles of each niche and disease sample. Pseudobulk profiles underwent similar filtering as described above, with niches containing fewer than nine samples excluded from analysis.

Distances between samples at each level were computed using the formula: *distance = 1 - corr(x, y)*, where *corr* represents the Pearson correlation coefficient between the vector values of sample *x* and sample *y*, generating a distance matrix. Each distance matrix was then used to generate a level-specific MDS plot. Then, distances were summed into a single matrix, retaining only samples with paired snRNA-seq and ST data to address missing values. This cumulative distance matrix was further used to perform joint MDS, summarizing differences across levels. Finally, silhouette coefficient, implemented in sklearn, was computed from each distance matrix at the sample level to quantify the clustering of lesion types.

3.13 Characterization of differences between lesion types

3.13.1 Differential expression analysis between lesion types

For snRNA-seq data, I performed pseudobulking per cell type and sample using the function get_pseudobulk from decoupler (min_cells > 10 and min_counts > 1000). Next, I filtered out lowly expressed genes for each cell type using the function filter_by_expr from decoupler (group = Lesion type, min_count > 10, and min_total_count > 15). I conducted differential expression analysis using PyDESeq2 (v.0.3.5), with lesion type and biological sex serving as design factors. I performed the Contrasts between different pairwise lesion types MS-CA vs CTRL, MS-CI vs CTRL, and MS-CA vs MS-CI (cooks_filter = False and independent_filter = False). I only performed the contrasts when a sufficient number of replicates were available for a particular cell type (min samples > 2).

For ST data,I performed pseudobulking per niche and slide for MS-CA and MS-CI typed MS slides. Similarly, I filtered lowly expressed genes, and conducted differential expression analysis per niche between the two lesion types using the same approach as described above.

3.13.2 Enrichment analysis of differential expressed genes

For snRNA-seq data, I filtered the genes based on significance criteria (adjusted P < 0,05, absolute (log2FC) > 1) and categorized them into lesion types according to the sign of the gene-level statistic. As genes could potentially be assigned to multiple lesion types due to the various

contrasts generated in the preceding section, I took measures to ensure uniqueness, resulting in a distinct list for each lesion type. Then, I conducted enrichment analysis for each cell type and contrast using the REACTOME gene set collection with the ORA method from decoupler. I excluded gene sets containing the terms FETAL, INFECTION, or SARS prior to enrichment analysis.

For ST data, I used all available gene-level contrast statistics between MS-CA and MS-CI per niche as input for the ULM method from decoupler to compute differential activity scores across niches using the same REACTOME gene set collection.

3.13.3 Compositional data analysis

Pau Badia i Mompel performed this subsection. Cell type compositions in snRNA-seq data were computed per sample by adding the number of cells per cell type, dividing by the total number of cells, and logarithmically transforming the obtained proportions using the clr function of the composition-stats python package. In this calculation, stromal cells and B cells were excluded to enhance comparability, as they were absent in most CTRL samples. Neurons were also excluded due to their presence being attributed to the nature of the tissue sampled rather than the lesion type. Cell subtype compositions in snRNA-seq data were computed per sample and cell type in the same manner.

Niche compositions in ST data were computed per sample by adding the number of spots per niche, dividing by the total number of spots in the slide, and logarithmically transforming the obtained proportions. In this calculation, gray matter and ependyma were removed due to tissue sampling reasons as described previously. Cell type compositions per niche and slide were computed by adding the deconvoluted cell type abundances per cell type across spots of each niche, dividing by the total sum of abundances for the niche, and logarithmically transforming the obtained proportions. Cell type compositions per slide were computed by adding all deconvoluted cell type abundances per cell type, dividing by the total sum of abundances across the entire slide, and logarithmically transforming the obtained proportions. For the latter two calculations, neurons were excluded as previously described.

Kruskal-Wallis tests were used to examine differences in each compositional type (adjusted $P < 0,05$). For significant elements, the Wilcoxon rank-sum test was employed to assess pairwise differences between lesion types (adjusted $P < 0,10$)

3.14 Differential cell-cell communication inference

The computational pipeline was designed by Pau Badia i Mompel, and I evaluated the biological results when discussing the design. Cell-cell communication inference involved combining results from both snRNA-seq and ST data to minimize false positive inferred interactions. Differential ligand-receptor interactions between the three lesion type contrasts across cell type pairs were inferred using the consensus ligand-receptor database from LIANA+ (v1.0.1). For each contrast, LIANA+ was utilized to compute a differential interaction score between a ligand of cell type A and a receptor of cell type B as the mean value between their differential gene statistics, generating lesion type-specific $cell_A$-$cell_B$-ligand-receptor interaction tetramers. Interactions with conflicting signs between ligand-receptor or with both genes not being significant (BH adjusted P < 0,15) were disregarded.

Due to the limited number of replicates for B cells, T cells, and Stroma cells for differential expression analysis, a separate strategy was employed to infer their communication events. For each contrast, snRNA-seq data was filtered for the lesion type being tested, retaining only genes that were expressed in at least 5% of cells, and the rank aggregate method from LIANA+ was utilized to infer cell-cell communication scores. Interactions involving any of the three cell types and deemed significant were retained (BH adjusted P < 0,15).

For the remaining significant interactions, spatially informed local scores were inferred across slides using a custom multivariate version of the normalized product method from LIANA+. In this approach, cell type proportions and gene expression were binarized for each slide (proportion > (1/Nº of cell types); log-normalized expression > 0), and feature values per spot were spatially weighted by averaging neighboring spots using an L1-norm Gaussian kernel (bandwidth = 150). To ensure comparability, features were normalized by their maximum value, bounding their values between 0 and 1.

The local score for a given $cell_A$-$cell_B$-ligand-receptor interaction was computed for each spot using the formula:

$$score = C_A * C_B * L * R,$$

where C_A and C_B are the normalized spatially weighted cell type proportions of cell type A and B, and L and R are the normalized spatially weighted gene expression values of the ligand and receptor genes.

Differences in interaction scores between lesion types were evaluated using the Wilcoxon rank-sum test. Significant interactions with no conflicting sign with the scores obtained in snRNA-seq were selected (BH adjusted $P < 0,15$). Candidate interactions were further filtered based on the results of cell subtype compositional changes and cell subtype marker genes described in the previous sections. Interactions were retained only if the ligand or the receptor was a marker gene for at least one cell subtype that significantly changed its abundance in the corresponding lesion type.

3.16 Illustrations

All illustrations present in the thesis were done by myself using the Affinity Designer (v1.10.8) vector persona.

3.17 Text editing

The text written in this thesis is my original work, and I used ChatGPT to improve its flow and readability.

4. Results

The results presented in the following four chapters refer to my original work, recently published[159]. I expand directly on these findings, providing additional context and deeper insights into specific areas.

4.1 Introduction to transcriptomics data

4.1.1 Selection and preparation of the MS cohort

My aim in the first part of the study was to select a cohort of subcortical white matter MS lesions to perform transcriptomic studies, gaining insight into key events driving lesion progression. Therefore, I began by histologically characterizing 21 fresh-frozen brain human tissue blocks, with the support of Annika Hoffman and Christian Riedl, and under the supervision of Simon Hametner and Lucas Schirmer.

To perform the histological analysis, I focused on three key spatial patterns of inflammation and tissue damage[9,10,160] identified from the IHC stainings I had previously performed (**Fig. 6a-b, Methods**). First, I assessed inflammation by detecting activated myeloid cells using anti-CD68 and anti-CD163 staining. This allowed me to identify acute (MS-A) and chronic active lesions (MS-CA), as chronic inactive (MS-CI) typically exhibit lower levels of myeloid cell activation. Secondly, I examined the presence of a demyelinated core using LFB and MOG stainings. The characteristic demyelinated hypocellular core, where the astroglial scar would form, allowed me to distinguish MS-CA and MS-CI lesions. Finally, in cases of MS-CA lesions, I also assessed the TBB staining to detect iron accumulation at the rim of lesions. Considering these characteristics and the possibility of multiple lesions within a single tissue block, I could characterize a total of 7 MS-A, 20 MS-CA and 12 MS-CI lesions (**Supplementary Table 1**).

To ensure that the samples were suitable for transcriptomic analysis, I measured the RNA integrity number (RIN) of all tissue blocks. The RIN is a measure of RNA quality, with higher values indicating better integrity. For this study, I selected only tissue blocks with a RIN value of 6 or higher (**Methods**). Many blocks did not meet the criteria and could not be included in sequencing or RNA-related analyses. Additionally, tissue availability was a challenge, as a substantial amount of tissue was required for both transcriptomic analysis and their subsequent validations. As a result, MS-A samples had to be excluded from the analysis.

After characterizing the lesions and addressing the challenges related to sample quality and availability, I selected the specific areas for transcriptomic analysis. It was crucial to ensure that the selected tissue areas closely matched the previous histological characterization to minimize the impact of lesion progression on the results. Because of that, I focused on choosing the areas that better represented different stages of lesion progression and made sure that the lesion rim was centered on the slides used for spatial transcriptomics (ST). Once I had obtained the tissue for ST, I continued to cut adjacent sections for single-nucleus RNA sequencing (snRNA-seq). This paired approach resulted in a final cohort comprising 6 CTRL, 8 MS-CA and 4 MS-CI lesions for ST (n=18), and 6 CTRL, 6 MS-CA and 4 MS-CI for snRNA-seq (n=16) (**Fig. 6c-d, Methods, Supplementary Table 2**).

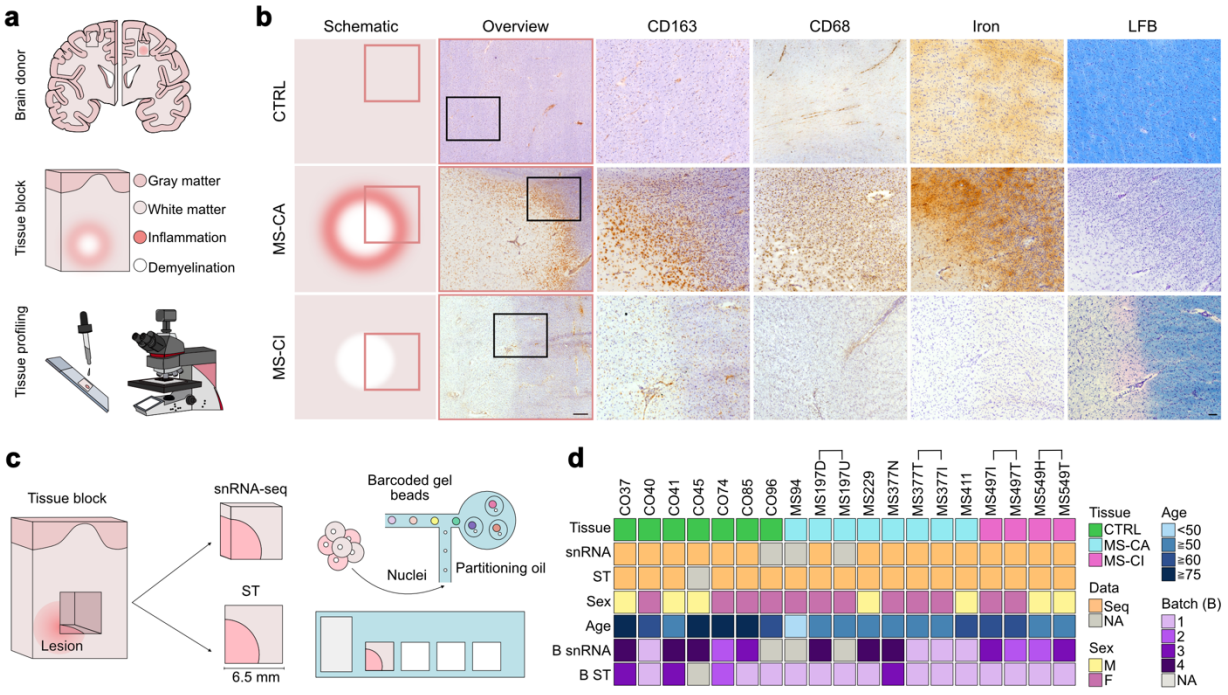


Figure 6. Lesion characterization and experimental design. Repurposed from Lerma-Martin *et al*. 2024. **a,** Schematic drawing of tissue preparation for lesion characterization. **b,** Histological assessment of MS lesions with IHC stainings for CD163 and CD68 (myeloid cells), iron or TBB (iron rim) and LFB (myelin). Overview scale bar 500µm. Zoom-in scale bar 100µm. **c,** Scheme of tissue collection for transcriptomics approach. **d,** Study design showing different data modalities and associated metadata. Links between samples indicate that they come from the same tissue block. B, Batch; NA, not available.

4.1.2 Cellular diversity in MS lesions

After selecting and obtaining the tissue from the samples, I isolated their nucleus, prepared the sequencing libraries and conducted quality control checks to ensure they met the necessary sequence requirements (**Methods**). I then sent the libraries for sequencing to obtain the FASTQ files. Following this, I developed scripts to process these files and generate the count matrices for data analyses (**Methods**).

I then performed the computational analysis, starting with the implementation of a strict quality control and the removal of low-quality profiles and potential doublets (**Methods**). Then, I generated an atlas that contained a total of 103.794 nucleus (n=16, average=6.487, mean genes per nucleus = 2.125). I proceeded to annotate this atlas using classical cell markers, and identified nine major cell types: oligodendrocytes (OL), oligodendrocyte progenitor cells (OPC), astrocytes (AS), myeloid cells (MC), endothelial cells (EC), T cells (TC), B cells (BC) and stromal cells (SC) (**Fig. 7a-b**). While my focus was on white matter MS lesions, when cutting the sections I also captured some surrounding gray matter (e.g., deep cortical layers), resulting in the presence of a neuron cluster (NEU).

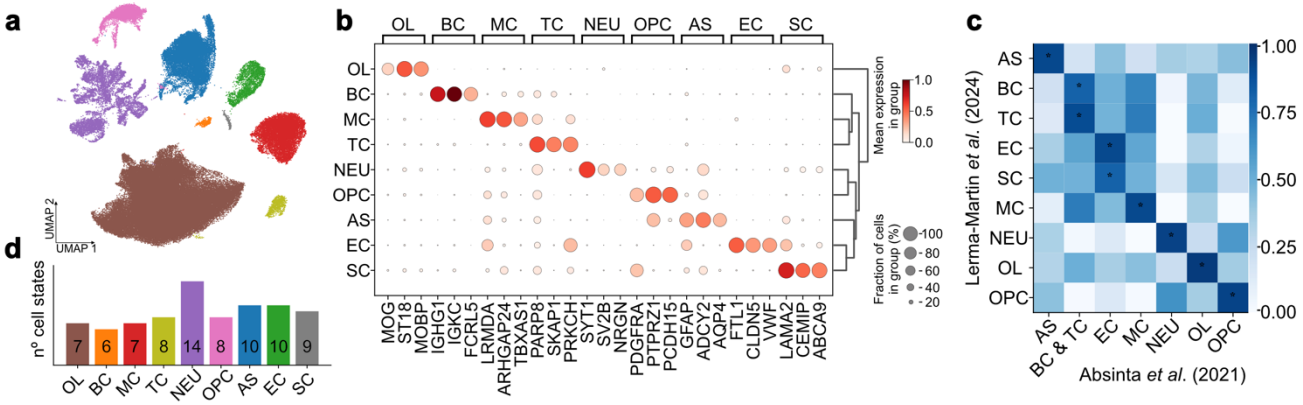


Figure 7. **Annotation of the MS snRNA-seq atlas.** Repurposed from Lerma-Martin *et al*. 2024 **a**, UMAP of integrated snRNA-seq data (n=103.794). **b,** Dotplot of averaged z-transformed gene expression of marker genes for each main cell type. **c**, Pearson correlation matrix of transcriptomic profiles between the snRNA-seq atlas of this study and Absinta et al.[114] (BH adjusted P < 0,05, r > 0,75). **d**, Barplot with the number of cell subtypes per main cell type.

To validate the annotation of these main cell types, I compared the atlas with another study focused on subcortical MS lesions[114], confirming the presence of all main cell types (**Fig. 7c, Methods**). Following this validation, I further enhanced the resolution of the dataset by sub-clustering all major cell types and annotating their cell subtypes based on existing literature. I identified 7 subtypes for OL, 8 for OPC, 10 for AS, 7 for MC, 10 for EC, 8 for TC, 6 for BC, 9 for SC and 14 for NEU, making a total of 79 unique cell subtypes (**Fig. 7d, Extended data Fig. 1a-h, Methods**).

4.1.3 Spatial mapping of MS lesions

Next, I performed ST to explore the spatial context of the lesions more deeply. I introduced this kind of technology into my laboratory using the 10x Genomics platform, Visium. As aforementioned, my goal when working with MS samples was to capture the lesion rim at the center of each capture area, allowing the investigation of spatial changes in between the peri-plaque white matter and the lesion core. To ensure precise and consistent tissue capture, I developed a novel scoring method using a 3D-printed rectangular prism with sharp edges (**Fig. 8a**), which allowed me to consistently capture the same tissue area for both transcriptomic techniques (**Methods**).

Since this was my first experience with spatial transcriptomics in human subcortical white matter, determining the optimal permeabilization time was essential (**Methods**). In this step, cells need to release an appropriate amount of mRNA for reverse transcription into cDNA. If too little mRNA is released, it may not yield enough for sequencing, while excessive release can lead to mRNA diffusion or cross-contamination between spots, resulting in false positives. Due to the unique characteristics of each tissue, permeabilization times vary; therefore, I tested durations from 5 to 30 minutes and found that 18 minutes was optimal for this tissue. With this information, I proceeded to perform the libraries and then sent the samples for sequencing to obtain their BCL files. I developed scripts to process these files, generating FASTQ files and count matrices for subsequent data analyses (**Methods**).

Next, I performed the data quality control, and obtained a total of 67.851 spots (n=18, mean spots per sample=3.769, mean of genes per spot=1.373) (**Methods**). Since ST slides can capture multiple cells per spot and lack single-cell resolution, I addressed this limitation by deconvoluting the slides. This approach involved inferring cell type composition across spots using the paired snRNA-seq atlas as a reference (**Fig. 8b-c, Methods**).

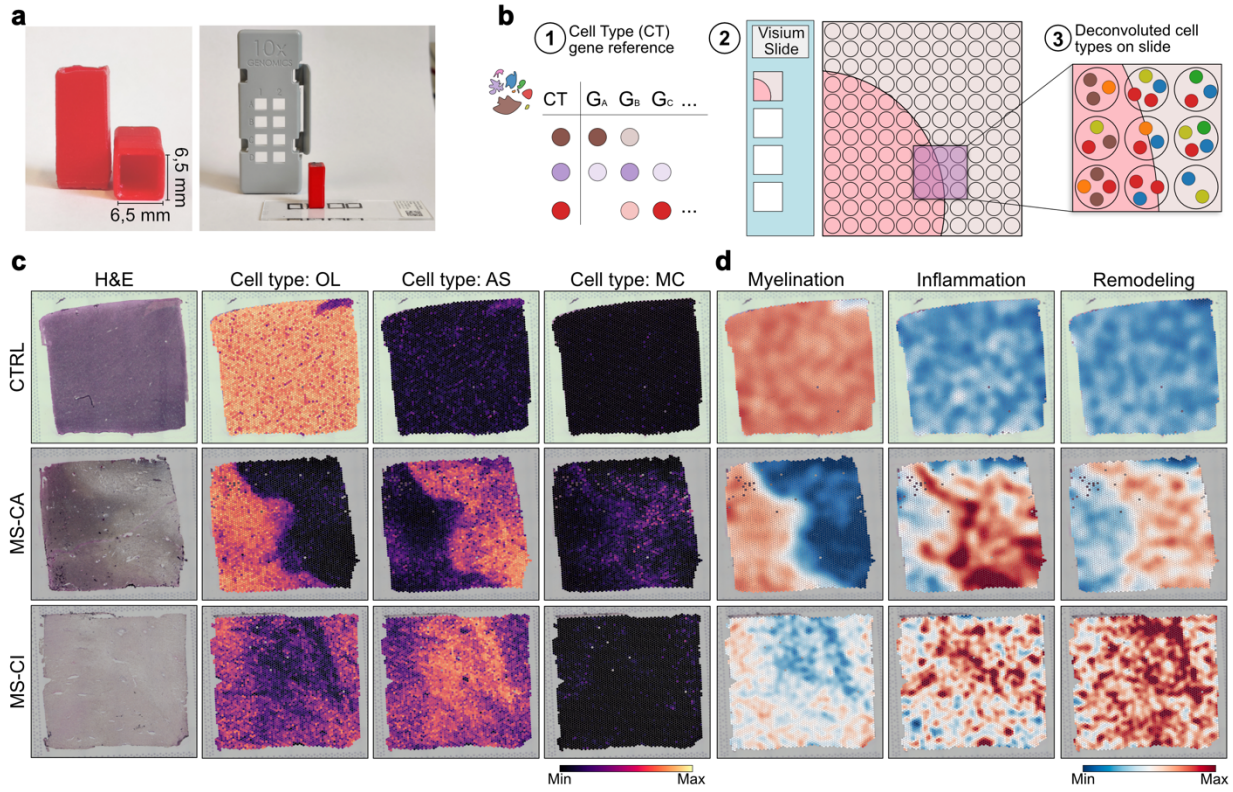


Figure 8. **Spatial characterization and deconvolution of MS lesions.** Repurposed from Lerma-Martin *et al*. 2024. **a,** Picture of the rectangular prism measurements and how its squared area fits within a Visium slide ROI. **b,** Scheme of deconvolution using (1) the cell type gene reference and (2) the spatial location of the lesion on the Visium slide, resulting in (3) distinct cell compositions within individual spatial spots. **c,** Spatial panel displaying hematoxylin and eosin (H&E) staining (left) and, feature plots of cell type deconvolution results for OL, AS and MC. **d,** Inferred pathway activities between CTRL and MS lesion types. Min, minimum; Max, maximum.

I confirmed that the mapped cell types aligned with the tissue architecture across CTRL and MS lesion types, thereby validating my approach. In CTRL samples, slides were predominantly composed by OL in the white matter, occasionally including gray matter. In MS lesions, OL were primarily located in the peri-plaque white matter, while AS were more prevalent in the lesion core, a characteristic sign of the astroglial scar[1]. MC cells were most abundant at the lesion rims and cores of MS-CA, with their presence diminishing and becoming scattered in MS-CI, with even lower presence observed in CTRL (**Fig. 8c, Methods**).

To characterize the slides and explore biological functions in lesion and non-lesion areas, I performed pathway enrichment analysis (**Methods**).I selected the top significant pathway for each condition and ensured these aligned with known biological functions specific to each condition. In CTRL and MS non-lesion areas, which have a high proportion of OLs, myelination activity was

enriched[161] (R-HSA-9619665). In MS-CA, pathways such as IFN-γ signaling (R-HSA-877300) were enriched. This pathway, when active in AS cells within the demyelinated core, has been reported to help mitigate inflammation[162], whereas in MC present at the inflamed lesion rims, might promote a shift towards a neurotoxic phenotype[163]. In MS-CI, pathways involved in tissue remodeling, such as assembly of collagen fibrils and other multimeric structures (R-HSA-2022090), were strongly mapped to demyelinated lesion areas characterized by a high presence of AS[99].

Collectively, the integration of snRNA-seq and ST allowed me to generate a comprehensive dataset. Using deconvolution prediction models, I was able to map the main cell types in both lesion and non-lesion tissue areas and linked them to different biologically relevant pathways.

4.1.4 Spatial characterization of MS niches

A significant challenge in the MS field lies in understanding how the cells transition from a stochastically distribution in the brain to assuming highly compartmentalized position during lesion formation[164]. To address this, I aimed to characterize the different cell communities present in both lesion and non-lesion areas. For that, I devised a robust approach that combines manual "biased" annotation with computational unsupervised methodology, enabling a comprehensive investigation of cell organization in subcortical MS tissue. The manual component of this approach involved a detailed histological and immunohistochemical assessment, complemented by the analysis of Loupe files generated from the Space Ranger counts pipeline (**Methods**). These files provided in-depth gene expression profiles of the spots from the ST slides. With these different kinds of data, Christian Riedl and myself curated a manual annotation of what we named areas, which included: gray matter (GM), white matter (WM), peri-plaque white matter (PPWM), lesion rim (LR), lesion core (LC), ependyma (EP) and claustrum (**Fig. 9a**).

While we tried to be precise, it is important to acknowledge that human-driven manual annotations are inherently biased and may introduce errors. To address this, Pau Badia i Mompel and I, designed an unsupervised approach using factor analysis[165]. This approach enabled the integration of gene expression data with deconvoluted cell type proportions, resulting in tissue region equivalents called niches (**Fig. 9b-c, Methods**). To evaluate the robustness of the niche annotations, we compared them to the manually annotated areas, revealing a high level of overlap between the two (**Fig. 9d, Methods)**. Additionally, we found that niches exhibited greater intra-niche similarity compared to inter-niche similarity in both gene expression and cell type proportions across slides (**Fig. 9e, Methods)**.

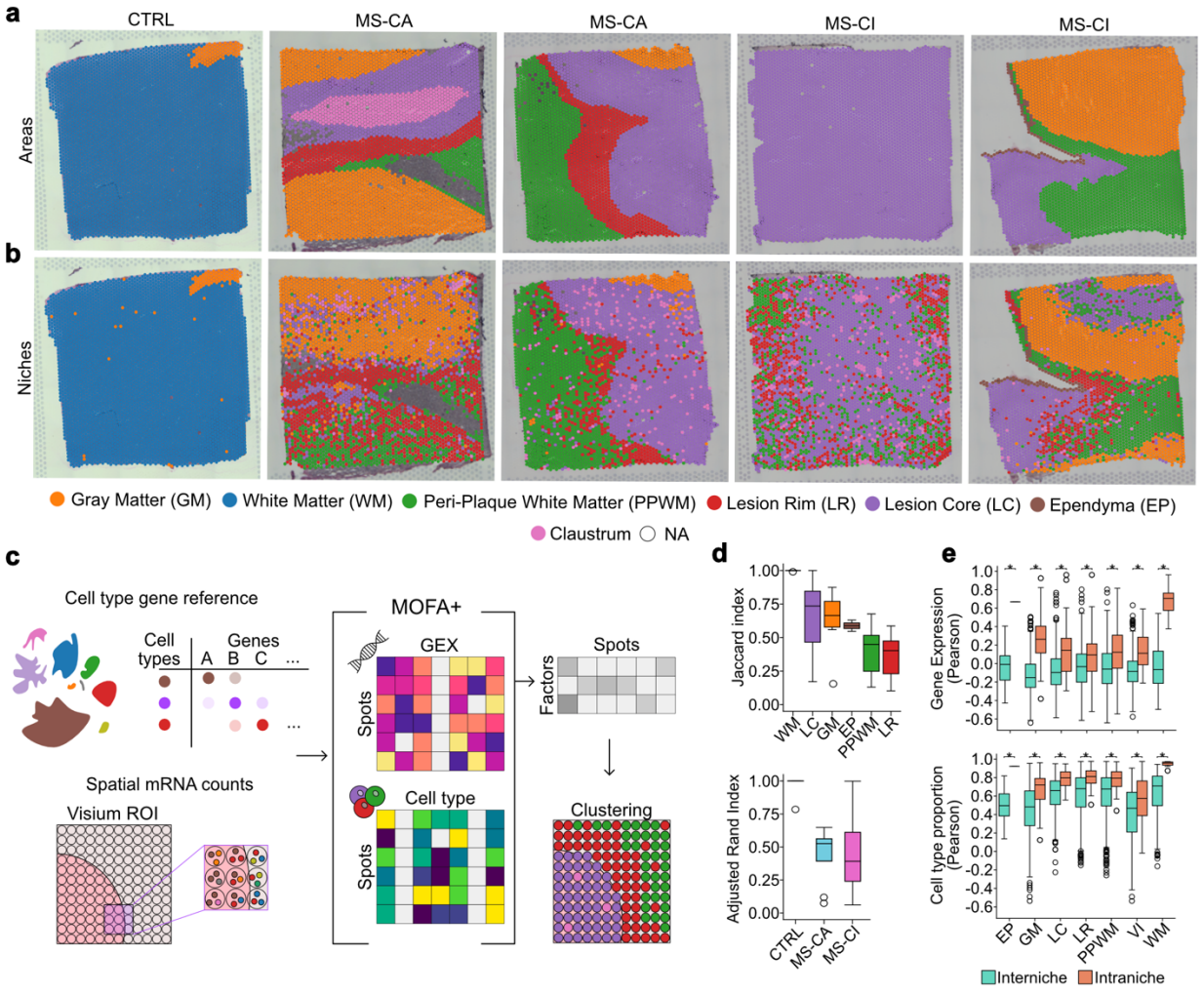


Figure 9. **Comparison of histopathological and computationally annotated spatial niches.** Repurposed from Lerma-Martin *et al*. 2024. **a,** Comparison of histopathological areas with **b,** the computationally annotated niches between lesion types. Two different MS-CA lesions are presented: one to showcase the sample containing the claustrum and the other as the representative MS-CA sample for the study. **c,** Pipeline design using MOFA+ to generate a spatial clustering from the integration of gene expression (GEX) and deconvoluted cell type proportions. **d**, Jaccard index between manually annotated areas and computationally annotated niches per shared category (top). ST spots belonging to annotation-specific categories were ignored in the calculation. Adjusted Rand index between histopathologically annotated areas and niches annotated in an unsupervised way per ST sample (bottom), values above 0 indicate agreement between annotations. **e**, Intra-niche and inter-niche Pearson correlations of normalized pseudobulked gene expression (up) and mean centered log-ratio transformed cell type proportion (bottom). Asterisks indicate significance (Wilcoxon-rank sum test, BH adjusted P < 0,05).

With this approach, we were able to predict the same areas that we previously annotated while also increasing the resolution of our analysis. This became evident in the precise mapping of the cell types within each niche, which showed high granularity across all tissue regions. I then characterized the distinct cell populations within the niches using markers that I had previously

used to characterize the main cell types (**Methods**). For instance, the main populations characterizing GM were NEU (*STY1*), while WM and PPWM were primarily composed of OL (*MBP, SOX10*). In inflamed LR, MC were prevalent (*CSF1R*), whereas scarred areas in the LC exhibited a higher concentration of AS (*GJA1*) (**Fig. 10a-b**).

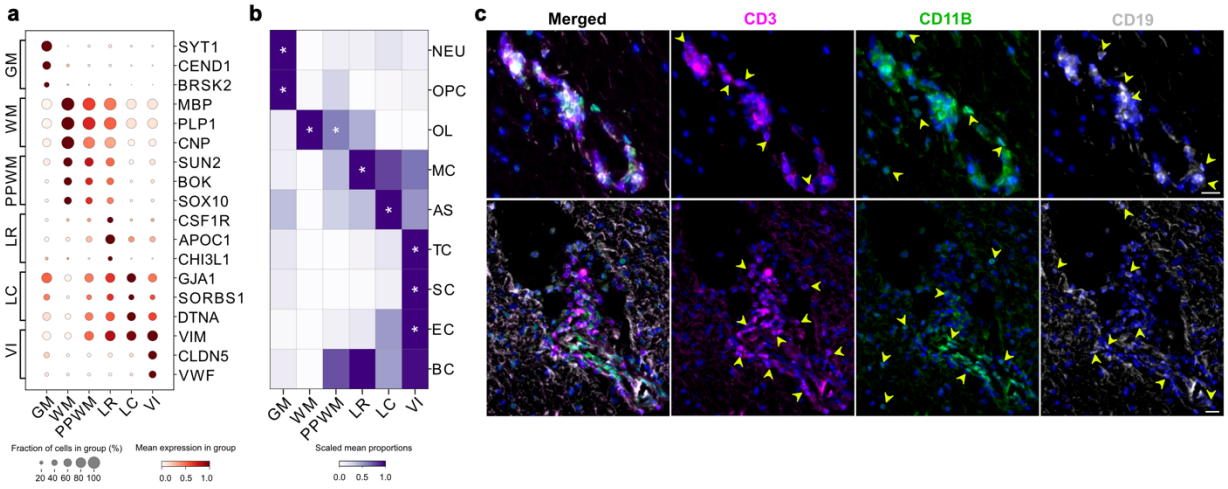


Figure 10. **Cell composition of the spatial niches.** Repurposed from Lerma-Martin *et al*. 2024. **a**, Dotplot of averaged z-transformed gene expression of marker genes for each niche. **b**, Heatmap of z-transformed cell type proportions for each niche. Asterisks indicate significance (adjusted P < 0,05). **c**, IF staining of blood vessels and perivascular spaces in MS-CA lesions with antibodies against CD3 (TC), CD11B (MC) and CD19 (BC). Scale bar 20µm.

Of note, I identified a previously unnoticed tissue niche, which I named vascular infiltrating (VI). This niche was characterized by perivascular spaces enriched with TC, SC and EC (*CLDN5*), and to a lesser extent, BC and MC (**Fig. 10a-b**). To validate the presence of this niche, I performed an IF staining on perivascular spaces in the lesion core and rim of MS lesions, which confirmed the infiltration of immune cells such as TC, MC and BC (**Fig. 10c, Methods**). Lastly, in reviewing the previously annotated areas, I decided to reclassify the claustrum as part of the GM, given that it is a thin, neuronal structure. Although I identified the EP, I opted not to include it in further analyses, as it did not contribute to my current focus. Overall, the niche annotation provided valuable insights into the spatial organization of cell communities within subcortical MS tissue.

4.1.4.1 Transcriptional changes in MS niches

To gain a deeper understanding of the different cell communities and their role in lesion progression, I examined the gene expression changes across different lesion stages : starting with the VI, then the LC, followed by the LR and finally PPWM and WM in control samples (**Methods**).

In the VI niche, genes such as *LAMA5* and *SOX18*, which are associated with EC, play crucial roles in maintaining endothelial integrity and control of angiogenic responses under inflammatory conditions[166,167]. In contrast, genes like *TUBB6, AMOTL1,* and *EMP2,* linked to pro-angiogenic tendencies, can potentially influence the tight junctions of the blood-brain barrier and facilitate vascular infiltration[168–170]. Additionally, genes such as *NECTIN2* and *CRIP1*, also enriched in the VI, suggest involvement in transendothelial migration of leukocytes[171] and activation of microglia and infiltrating macrophages[172], respectively, contributing to the inflammatory milieu (**Fig. 11a**). Moreover, I explored enriched pathways in the VI niche, identifying TGF-β as pivotal in tissue remodeling and the JAK-STAT pathway linked to tissue inflammation, in line with the genes that I previously described (**Fig. 11b, Methods**).

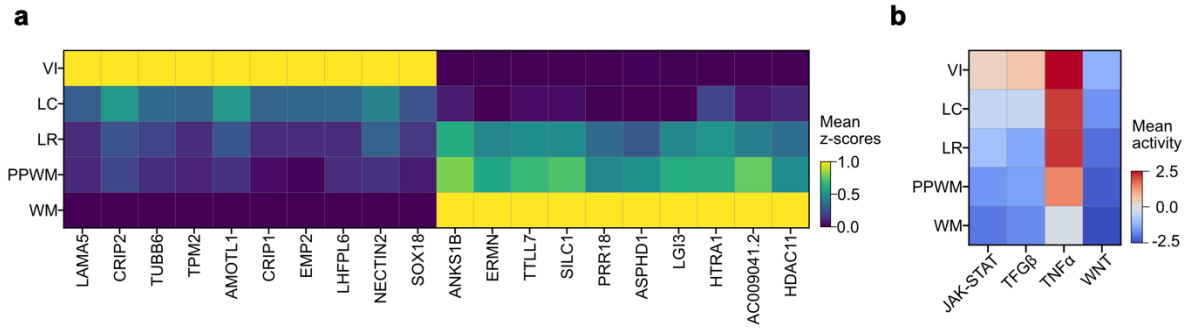


Figure 11. Spatial trajectories and pathway activity in niche progression. Repurposed from Lerma-Martin *et al*. 2024. **a**, Heatmap of z-scaled expression of top 10 genes (positive and negative scores) per niche based on Spearman correlation corresponding to a spatial trajectory from MS lesion to CTRL white matter (VI-LC-LR-PPWM-WM). **b**, Mean pathway activity across ST samples grouped by niche.

As the lesion and inflammation progress through MS-associated niches, they eventually reach the PPWM niche, located at the periphery. While this niche histologically resembles control WM, it exhibits enrichment of inflammatory signaling pathways, such as TNFα, and expresses genes like *ASPHD1*, previously linked with MS risk[173]. However, due to its high OL population, the PPWM also shares characteristics with control WM, expressing genes associated with myelin maintenance, its structural integrity, and the regulation of homeostatic OL function. For instance, *HDAC11* encodes for a protein that modulates myelin function[174], *ERMIN* determines myelin cytoskeletal organization and structure[175], and *LGI3,* along with *ANKS1B*, contribute to the modulation of synaptic plasticity in myelinated axons[176,177]. These findings suggest that the unsupervised characterization of the tissue niches mirrors the organization within subcortical MS lesion and non-lesion areas.

4.2 Novel identification of astrocyte cell subtype

4.2.1 Identification of ciliated astrocytes

In this chapter, I aimed to characterize the molecular diversity of astrocytes (AS) within subcortical MS lesions, as this major cell type can play a critical role in either propagating inflammation or fostering tissue homeostasis[178,179]. From the main atlas, I subclustered the AS population and characterized a total of 10 distinct subtypes based on existing literature: gray matter AS (GM), homeostatic white matter AS (Homeo), stressed AS (Stress), transitioning ciliated AS (TransC), ciliated AS (Cilia), reactive AS (React), phagocytic AS (Phago) and three different disease associated AS (Dis1, Dis2, Dis3) (**Fig. 12a-b, Methods**).

Next, I validated my annotations by conducting a comparative analysis with a previously published snRNA-seq atlas[114], which confirmed the presence of most subtypes based on their gene expression patterns (**Fig. 12c, Methods**). Interestingly, the largest number of shared genes was found between the Cilia AS population and the Cluster 12 from the other study, which they had labeled as senescent. Cellular senescence is an ubiquitous process that involves cell-cycle arrest and release of inflammatory cytokines. While this process is normally regulated within physiological and homeostatic processes, it can also occur in disease[185].

In reviewing the literature on ciliated AS I found no publications that specifically described their presence in white matter. While some atlases did include ciliated AS[173,180], they were not characterized and were typically grouped within the general AS population rather than reported as a distinct subtype. Therefore, I aimed to determine whether it was indeed a novel cell subtype or if it had already been identified and classified as senescent.

I began by assessing the composition of each sample at the cell subtype level (**Methods**). I noted that the Cilia AS had an overrepresentation of two samples that contained the EP niche (**Fig. 13a-b**). Given that both astrocytes and ependymal cells stem from radial glia, they share a similar molecular profile[181]. To prevent a potential characterization bias, I excluded these samples from the analysis.

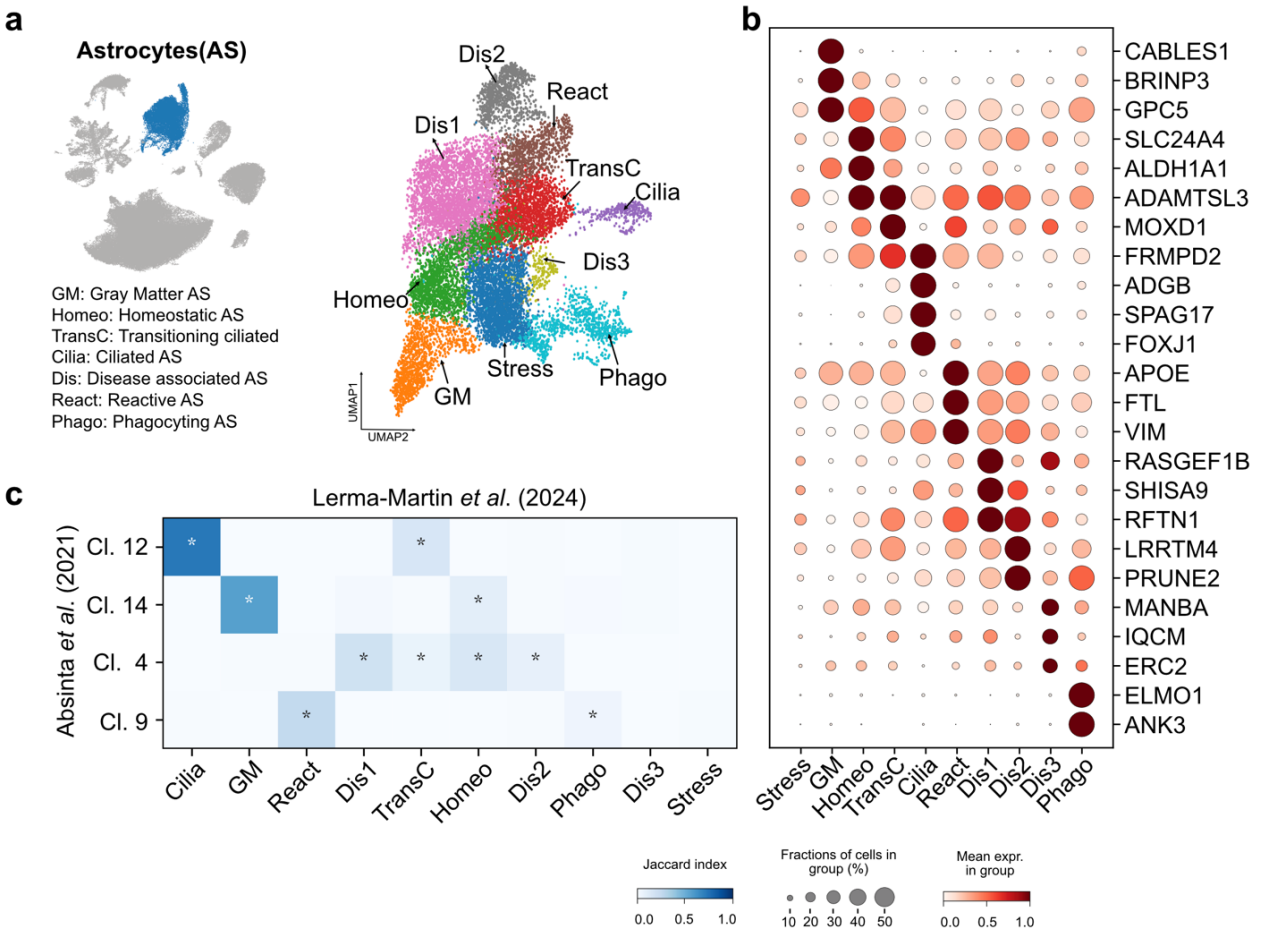


Figure 12. **AS subtype characterization and identification of ciliated AS.** Repurposed from Lerma-Martin *et al*. 2024. **a**, UMAP of AS subtypes based on snRNA-seq. Color indicates assigned AS subtypes. **b**, Dotplot of averaged z-transformed gene expression of marker genes for each AS subtype. **c**, Heatmap of Jaccard indexes between the top 100 marker genes ranked by corrected p-values (Benjamini-Hochberg) of the cell subtypes in this study compared to the subclusters described in the study by Absinta *et al.*,[114]. Cl, Cluster.

Then, I aimed to identify the unique gene expression patterns of this cell subtype and evaluated their possible functional roles based on the reported gene functions (**Fig. 13c, Methods**). The top expressed genes included *DNAH11* and *DNAH6*, two dynein family members crucial for the function and structure of the motile cilia axoneme[182], as well as *CFAP54* and *CFAP299,* which encode proteins associated with cilia and flagella[183]. Additionally, *SPAG17*, an important microtubule-stabilizing protein of the axoneme[184], was also present. These genes were also found to be upregulated in the analogous senescent AS cluster[114](**Methods**). Therefore, I tested for *SPAG17* and *CFAP299* in both datasets and observed that there was a significance difference in expression within the Cilia AS population, as expected. In contrast, analysis of classic senescent markers *CDKN2A* and *CDKN1A* revealed no significant differences in either atlas (**Fig. 13d**).

To further characterize this population, I conducted an enrichment analysis using a curated senescence gene set from a recent publication[158] as well as a dataset derived from the top 100 genes defining the ciliated population (**Methods**). Consistently, I observed significant differences when analyzing the cilia score across both datasets, whereas my analysis revealed no significant differences in senescence scores between the ciliated population and the remaining AS population in either dataset (**Fig. 13e**). Finally, I examined the proportion of senescent cells, finding that the overall AS population exhibited a higher percentage of senescent cells compared to the Ciliated AS (**Fig. 13f**). These findings support that the characterized population is indeed ciliated AS.

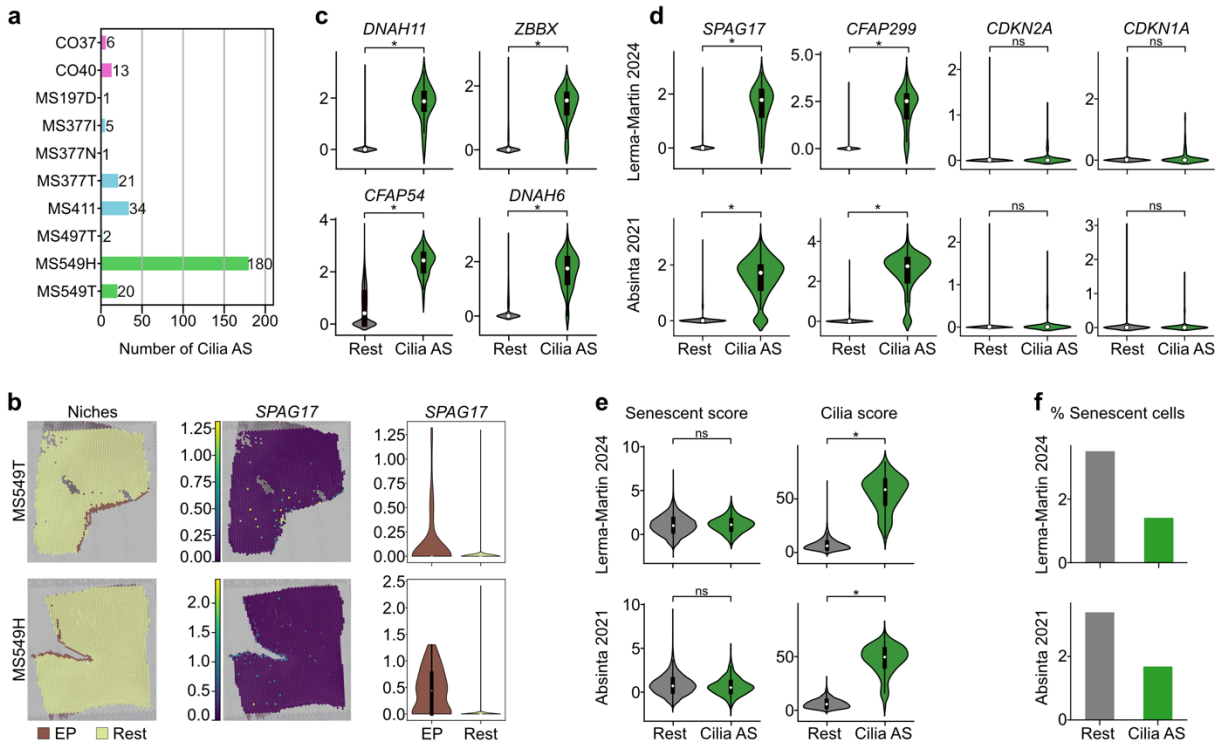


Figure 13. **Comparison of ciliated and senescent AS populations.** Repurposed from Lerma-Martin *et al*. 2024. **a**, Bar plot shows number of Cilia AS cells per sample. Color indicates condition. **b**, ST feature and violin plots for *SPAG17* for the two samples containing EP niches: MS549H and MS549T. **c**, Violin plots of representative Cilia AS marker genes compared with other AS subtypes. **d,** Violin plots of representative Cilia AS marker genes and senescent genes genes in AS population (rest) vs ciliated AS for both atlases. **e**, Violin plots of enrichment scores for senescent genes in AS population (rest) vs ciliated AS for both atlases. **f**, Barplot with the proportion of senescent cells in AS population (rest) vs Cilia AS for both atlases. Asterisks indicate significance (t-test, BH adjusted P < 0,05).

4.2.2 Spatial validation of ciliated astrocytes

After the initial characterization, I sought to gain insight into the spatial distribution of this AS cell subtype within the MS lesions. To achieve this, I predicted the localization of their expression profiles on the ST slides, observing a strong localization within the demyelinated core, particularly within the LC niche (**Fig. 14a**). To validate this prediction, I used single molecule fluorescence *in situ* hybridization (smFISH) as it directly targets individual mRNA molecules, providing a precise confirmation of the transcriptomic data (**Methods**).

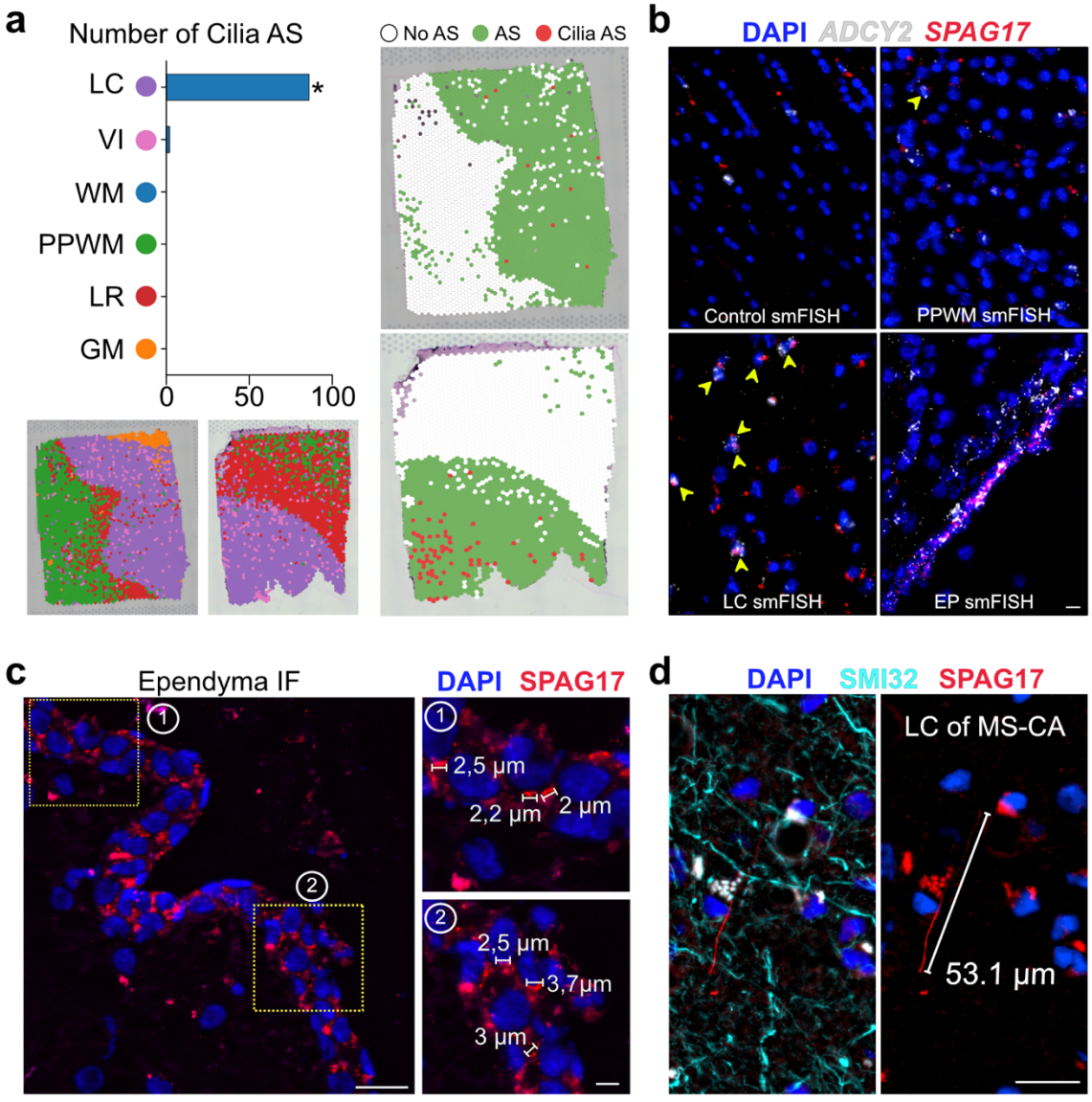


Figure 14. Extended legend in next page

Figure 14. **Spatial localization and initial validation of ciliated astrocytes.** Repurposed from Lerma-Martin *et al*. 2024. **a,** Predicted number of Cilia AS per niche. Asterisk indicates significance (empirical P < 0.05) (top left). Spatial feature plots showing the niches (bottom left) and spots predicted to contain Cilia AS (right). Color indicates absence of AS (white), presence of AS (green) and presence of ciliated AS (red). **b**, smFISH for *SPAG17* and *ADCY2*. **c,** IF staining of cilia (SPAG17) in ependymal cells adjacent to the lateral ventricle. **d,** IF staining of damaged axons (SMI32) and cilia (SPAG17) in the lesion core (LC) of MS-CA lesions. Length of cilia given in µm. Scale bars: 20 µm.

I used probes *ADCY2* to identify AS and *SPAG17* to confirm the presence of cilia. Indeed, I observed an increased number of *ADCY2*-*SPAG17* co-expressing cells within the LC of MS-CA, with smFISH signals comparable to those found in ependymal cells (**Fig. 14b**). To confirm that the mRNAs were translated into protein, I performed IF stainings using anti-SPAG17 antibodies (**Fig. 14c-d, Methods**). Before starting, I wanted to verify that the antibody was specific for cilia, so I tested it on ependymal cells. I examined samples that contained ventricles and observed ciliated structures in their inner lining, with an average length of 2,5µm, proving that it effectively stained cilia (**Fig. 14c**).

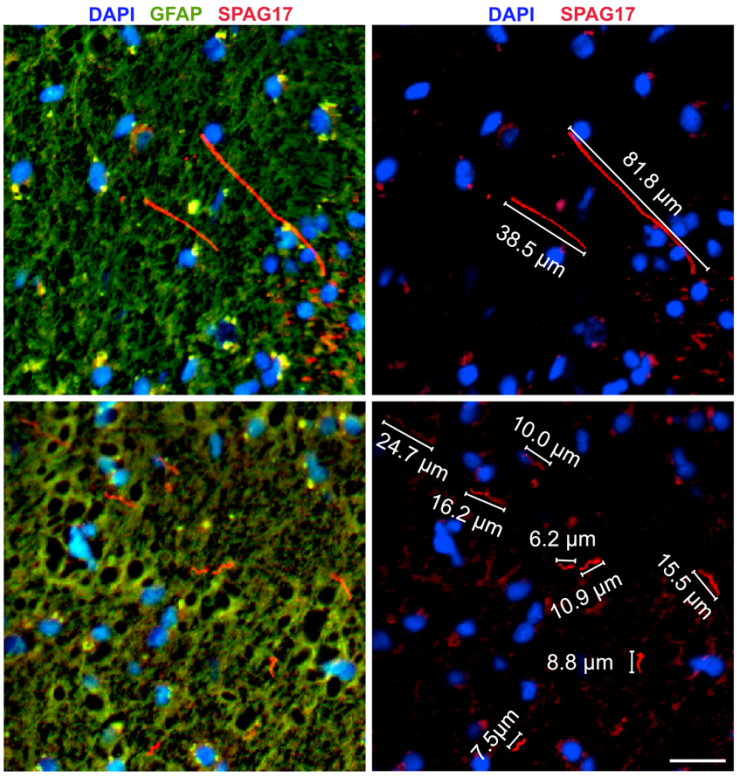


Figure 15. Variability of ciliated AS lengths at the LC of chronic MS lesions. Repurposed from Lerma-Martin *et al*. 2024. Two IF examples showcasing the length of AS cilia in µm in the LC of MS-CA lesions. Scale bars: 20 µm.

Since the predicted location of Cilia AS was within the lesion core, I wanted to ensure that the observed signal that I would get originated from cilia rather than from damaged axonal fibers. To do this, I assessed the overlap between SPAG17 and the neurofilament-specific SMI32 signal, an indicator of axonal damage, within the demyelinated core, and I could confirm the absence of overlap between the structures (**Fig. 14d**). With the preliminary validations completed, I was confident in the specificity of the signal and I could identify numerous cilia at the core of MS-CA lesions (**Fig. 15**). They varied in length, with some reaching to 82µm, a striking contrast with the length observed in ependymal cells.

4.2.3 Molecular characterization of ciliated astrocytes

After confirming the presence of ciliated astrocytes, I proceeded to investigate their molecular function by conducting an enrichment analysis of associated marker genes. This analysis primarily highlighted pathways closely linked to cilia biogenesis, suggesting a specialized role in the generation of these structures (**Fig. 16a, Methods**). Additionally, I conducted an analysis to detect enrichment of activated transcription factors within this population. This revealed a significant increase in activity of three transcription factors that were enriched in the Cilia AS subtype when compared to the rest: FOXJ1, REST, and TBX1 (**Fig. 16b, Methods**).

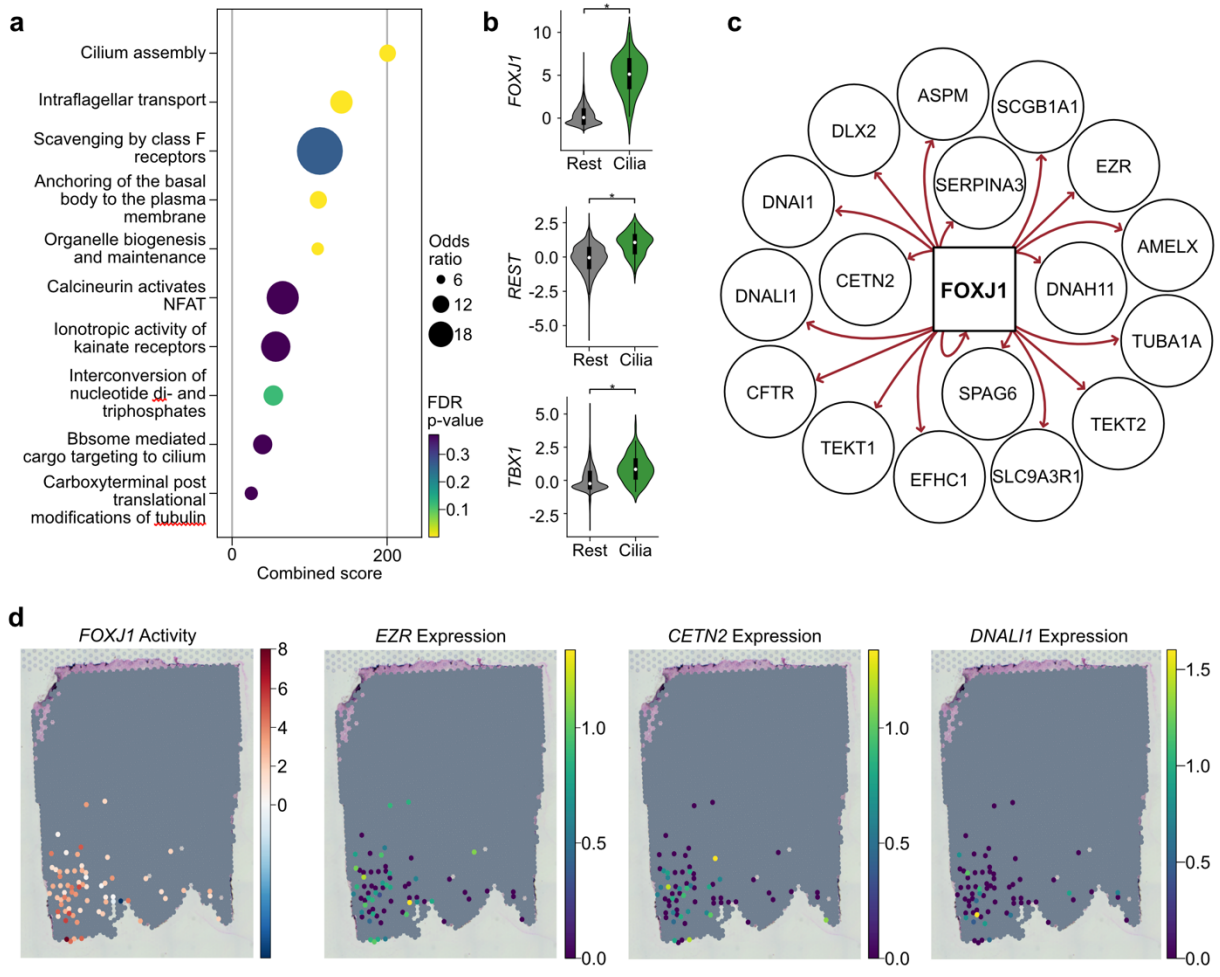


Figure 16. **Characterization of enriched pathways and transcription factors in Cilia AS**. Repurposed from Lerma-Martin *et al*. 2024. **a**, Dotplot of top ranked enriched pathways. Color indicates overrepresentation significance by Fisher exact test, and size indicates odds ratio between Cilia AS marker genes and each gene set. **b**, Violin plots of predicted transcription factor activity from snRNA-seq. Asterisks indicate significance (two-tailed t-test, BH adjusted P < 0,05). **c**, Transcription factor FOXJ1 (square) and its target genes (circles). **d,** Spatial feature plots showing FOXJ1 predicted transcription factor activity (blue/red) and gene expression of some of its target genes (purple/green).

FOXJ1, in particular, is a transcription factor crucial for the formation of motile cilia. To investigate its role in this subpopulation, I examined the target genes it regulates. This analysis revealed that FOXJ1 also regulated the expression of other essential genes that encode for ciliary components, such as *EZR* and *CETN2* [186] (**Fig. 16c**). These genes, along with FOXJ1 activity, were enriched in the same spots where ciliated AS cells were predicted to be localized, further supporting their functional significance within this population (**Fig. 16d**).

In summary, by integrating transcriptomics with various validation techniques, I confirmed the existence of a specialized AS subtype within the glial scar of MS lesions, characterized by a high degree of specialization towards generating large ciliated structures.

4.3 Dissecting MS lesion type differences

In this chapter I aim to gain insight into lesion progression, a task that has its own set of challenges. Lesion characterization typically relies on histological assessment, which, while informative, does not fully capture the underlying transcriptomic differences among the lesions. To address this, I collaborated with Pau Badia i Mompel to apply a multidimensional scaling (MDS), a technique useful for visualizing and interpreting the molecular and compositional variability in a 2D low-dimensional space within our pair transcriptomic dataset across samples (**Methods**). This analysis revealed that samples clustered based on their lesion type, a pattern that suggests greater differences between patient groups rather than variations within groups (**Fig. 17a**).

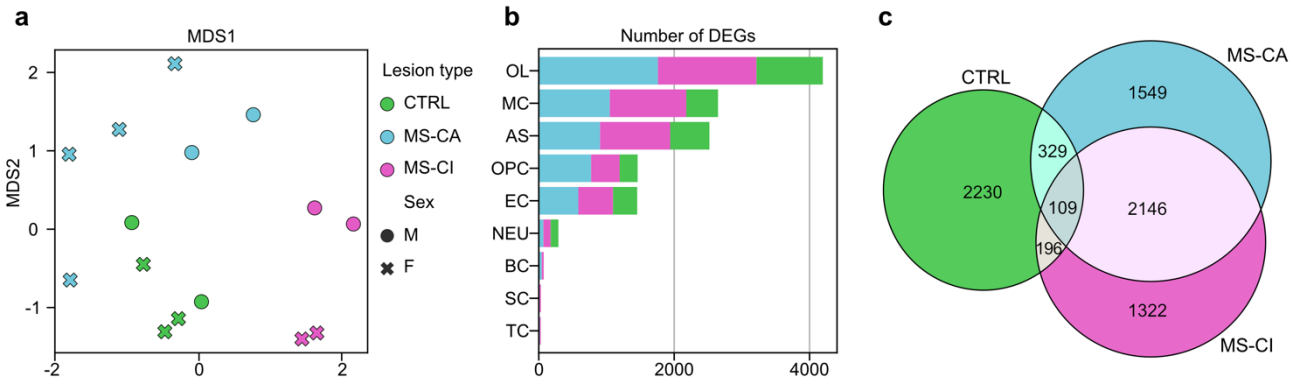


Figure 17. **Molecular and compositional differences across control and MS lesion types.** Repurposed from Lerma-Martin *et al*. 2024. **a**, Multidimensional scaling (MDS) plot based on the aggregation of molecular and compositional differences across paired samples. Color indicates MS lesion type. Shape indicates biological sex. **b**, Cumulative number of differentially expressed genes (DEGs) per cell type across conditions. Color indicates conditions. **c**, Venn diagram visualizing overlap of aggregated DEGs between CTRL and MS lesion types.

Having identified unique transcriptomic patterns for each lesion, I next sought to determine whether specific genes were associated with each condition. Using the snRNA-seq atlasidentified differentially expressed genes (DEGs) by comparing control samples to both types of MS lesions (**Fig. 17b, Methods**). The majority of dysregulated genes were specific to MS, with each lesion displaying its own distinct signature, consistent with my earlier findings. Additionally, a core set of genes was shared between MS-CA and MS-CI, suggesting potential similarities between both lesions. In contrast, CTRL tissue exhibited a unique signature, with a small overlap of genes shared with both MS-CA and MS-CI (**Fig. 17c**). These results suggest a clear transcriptomic signature and unique gene expression profile for MS lesions.

4.3.1 Cellular adaptations in MS progression

I continued my analysis by focusing on the differentially expressed genes within the main cell types that showed the most significant changes between CTRL and MS lesion types: OL, MC, AS, OPC, and EC (**Fig. 17b, Methods**). From these cell types, together with Pau Badia i Mompel, I then explored whether their respective cell subtypes experienced any changes in abundance across disease conditions, which could potentially help explain the different transcriptomic patterns. Interestingly, only OL, MC and AS cell subtypes showed significant changes in abundance, with OPC cell subtypes also exhibiting alterations, but to a lesser extent (**Methods**). Therefore, I proceeded to conduct a more in-depth characterization of these three main cell types to better understand their adaptation from a healthy to a pathological microenvironment.

4.3.1.1 Oligodendrocytes

In oligodendrocytes (OL), distinct gene expression patterns emerge between CTRL and MS, reflecting the different cellular function and response to disease. These changes can be directly correlated to the variations in abundance of some of its cell subtypes (**Fig. 18a, Extended Data Fig. 1a**). Genes enriched in CTRL mostly came from Homeo1 and Homeo3, with a lesser contribution from Homeo2. These genes are linked to classical OL functions, such as myelin production, maintenance and stability (*ADAMTS4, EPHB, ELOVL6*) as well as differentiation (*ERBB2, NDE1, CDK18*) (**Fig. 18a**).

However, OL phenotype in both MS lesions was influenced by the cell subtypes Dis1 and Dis2. In MS-CA lesions, enriched genes are related to inflammation (*EIF5, NFKB2, IRF)*, general cell and ER stress responses (*ATF4, HSPB1, HSP90B1*) and antigen presentation (*CD274*), aligning with the findings from the pathway enrichment analysis (**Fig. 18a, d**).

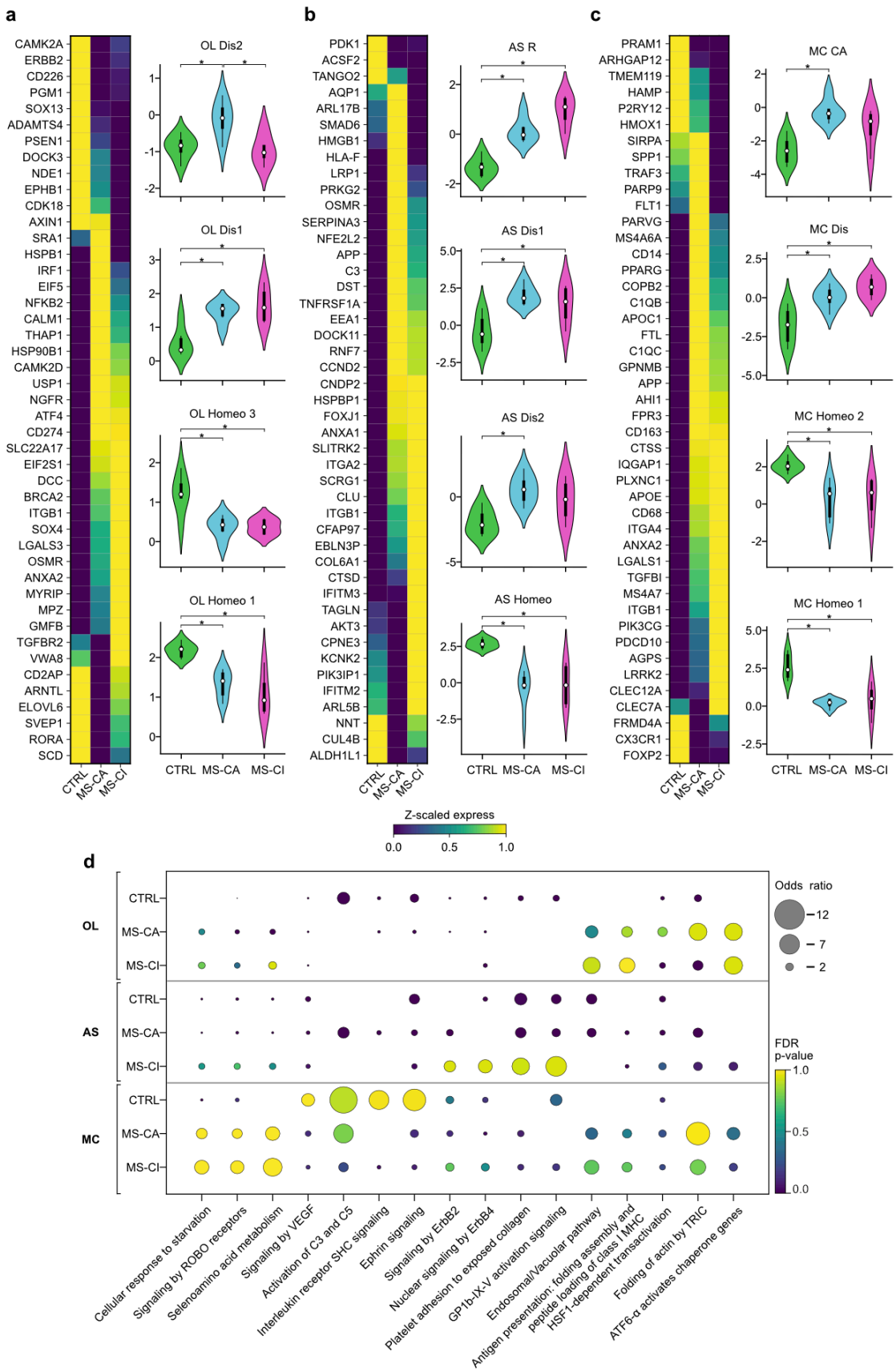


Figure 18. Extended legend in next page

Figure 18. **Changes in OL, AS and MC across MS lesions**. Repurposed from *Lerma-Martin et al*. **a**, OL DEGs (left) and subtype composition (right) associated with DEGs. **b**, AS DEGs (left) and subtype composition (right) associated with DEGs. **c**, MC DEGs (left) and subtype composition (right) associated with DEGs. **d**, Dotplot of enriched pathways for DEGs of OL, AS and MC, grouped by condition.

In MS-CI lesions, there are genes associated with cell death (*DCC*) and severe stress (*OSMR*, *SLC22A17, BRCA2*), indicating a basal inflammation that is supported by the pathway analysis results (**Fig. 18a, d**). In addition, there is also presence of genes related to counteracting inflammation and tissue remodeling activity (*TGFBR2*), lipid and myelin regulation (*LGALS3, MYRIP, MPZ, NGFR*) and cell differentiation (*SOX4, GMFB*) (**Fig. 18a**).

4.3.1.2 Astrocytes

Astrocytes (AS) exhibit a diverse range of phenotypes in MS (**Fig. 12a-b**), indicating distinct gene expression patterns between CTRL and MS lesions. In CTRL, AS phenotype was associated with the Homeo cell subtype. This condition is characterized by the enrichment of genes involved in maintenance of homeostasis (*TANGO2, ALDH1L1, NNT*) and differentiation (*PDK1*) (**Fig. 18b**).

Interestingly, the gene expression profile in MS lesions correlates with the increased abundance of React and Dis1 and Dis2 AS cell subtypes. In MS-CA lesions, genes are implicated in various functions such as inflammation (*HMGB1, HLA-F, SERPINA3, C3, TNFRSF1A, OSMR*), demyelination (*PRKG2*), cell survival by inhibition of apoptosis (*RNF7*) and modulation of cell flow in blood vessels (*SMAD6*). Conversely, the presence of genes associated with anti-oxidative mechanisms (*NFE2L2*) and debris degradation (APP, LRP1, *EEA1*) suggests a complex interplay of responses within the MS-CA microenvironment. Of note, there is a transitional area with overlapping signals between lesions, characterized by genes involved in inflammation (*ANXA1, CLU, ITGB1*), cell stress (*HSPBP1*), axon guidance (*SLITRK2*) and cilia motility functions (*FOXJ1, CFAP97*) (**Fig. 18b**).

For MS-CI lesions, I observed a distinct set of genes related to proliferation (*CPNE3*), anti-inflammation (*TAGLN, AKT3*) and clearance of debris (*CTSD*). Interestingly, upregulation of antiviral genes (*ARL5B, IFITM3, IFITM2*) related to preventing their entry and propagation, suggests a potential role of AS in combating viral infections. Moreover, the presence of genes involved in scar formation (*ITGA2, COL6A1*) and the pathways associated with blood-brain barrier integrity and vessel remodeling, points to a broader involvement of AS in tissue repair and vascular dynamics in MS pathology (**Fig. 18b, d**).

4.3.1.3 Myeloid cells

Myeloid cells (MC) play key roles in MS, exhibiting distinct patterns between CTRL and MS lesions, contributing to the wide phenotypic range observed in disease (**Extended Data Fig. 1c**). In CTRL, two of its cell subtypes undergo changes in abundance, with only Homeo contributing to the enrichment of genes associated with this condition. The primary function of these genes is to maintain tissue homeostasis and regulate physiological MC functions (*P2RY12, CX3CR1, FRMD4A*). Of note, other genes present (*ARHGAP12*) encode for proteins involved in synaptic maintenance, which is consistent with pathway enrichment analysis results (**Fig. 18c, d**).

In disease, CA and Dis cell subtypes were associated with the evolving phenotype of MC through disease. In MS-CA lesions, genes related to MC activation (*TRAF3, PARP9*) and inflammation (*SPP1, APP, FTL, GPNMB*) are enriched, alongside an increase in complement activation (*C1QB*, *C1QC*). Interestingly, a simultaneous enrichment of anti-inflammatory genes (*MS4A6A, APP, PPARG*) suggest a dynamic response to the inflammatory milieu. In addition, there is an active cytoskeleton remodeling observed both at the gene level (*PARVG, IQGAP1*) and active pathways, reflecting an adaptability to the environmental cues and the plasticity of MC cells (**Fig. 18c, d**).

The transition from MS-CA to MS-CI lesions presents a complex microenvironment, characterized by the shared inflammatory (*APOE, CD68*) and anti-inflammatory (*CD163*) gene expression patterns. In MS-CI lesions, MC exhibit genes related to pattern recognition, cell motility and cell-cell signaling, that under this milieu are also indicators of late ongoing inflammation (*ANXA2, LRRK2, CLEC12A, CLEC7A, ITGB1*), also reflected in several enriched pathways. However, there are also genes involved in restoration of the homeostasis (*TGFB1, AGPS*) and proliferation (*PIK3CG*), indicating a shift towards normalcy after an acute inflammation (**Fig. 18c**).

Collectively, these results offer a comprehensive understanding through gene expression alteration of three main cell types that are key for lesion progression in MS.

4.3.2 Niches adaptations in MS progression

In these previous sections, I have characterized the progression of the lesions from a broad perspective to individual cell types. However, what truly drives the disease is the relationship that these cells have with their neighbors and surroundings. Therefore, for this next analysis, I aimed to delve deeper into the differential gene expression between tissue niches within the MS lesion types (**Methods**).

To investigate the most affected conditions and niches, I first focused on identifying those with the highest levels of gene dysregulation. This analysis revealed that MS-CA lesions exhibited the highest number of gene changes, particularly within the LR and LC niches (**Fig. 19a**).

Building on this, I examined pathway activity changes within the lesions to determine if they were associated with specific niches (**Methods**). In MS-CA lesions, immune system pathways were active across all niches, while MS-CI lesions showed pathways related to stress, like cell starvation, and remyelination (**Fig. 19b**). Surprisingly, the TGFβ pathway signaling, which is crucial for maintaining the homeostasis, was enriched in all niches except VI (**Fig. 19b**).

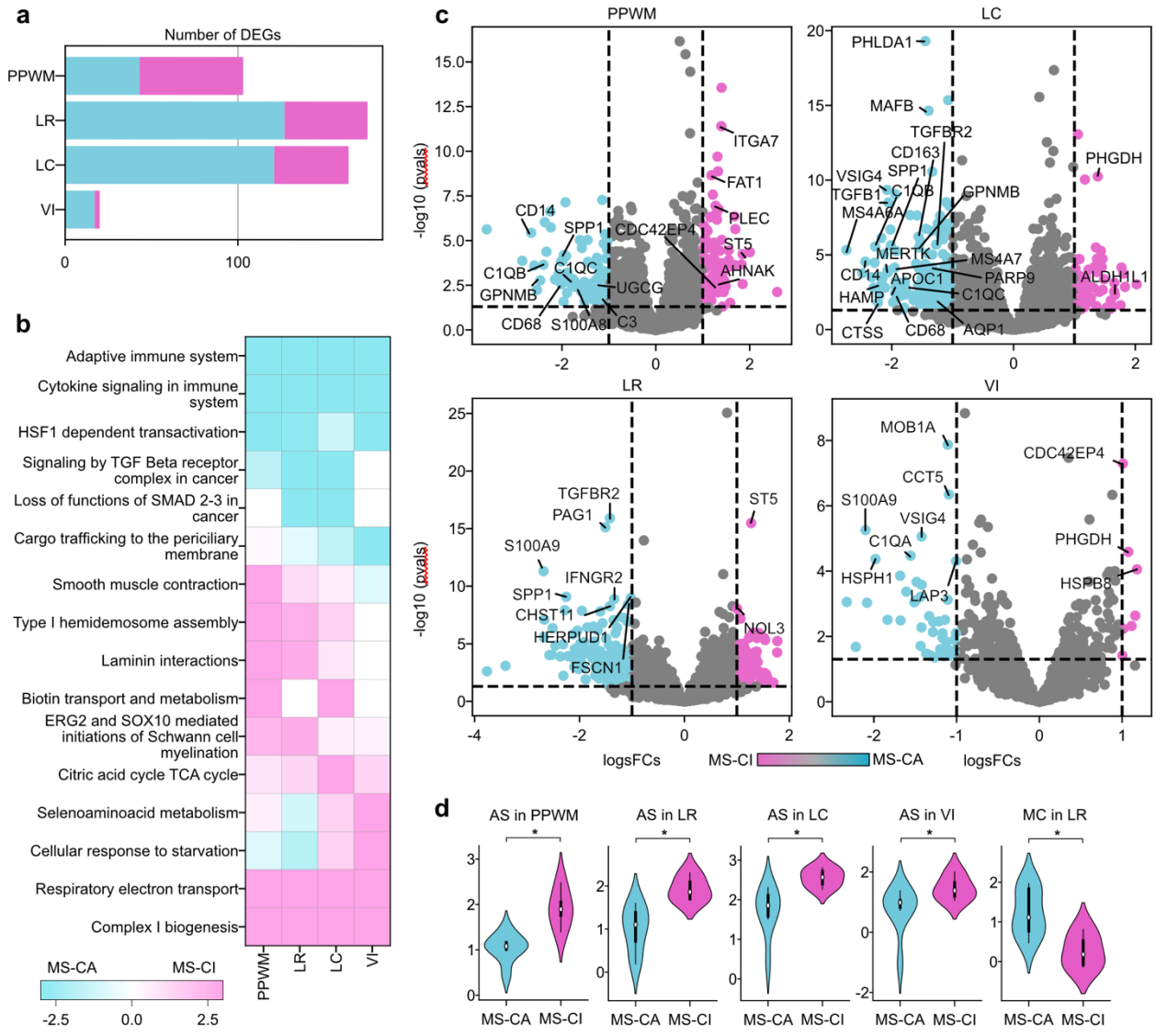


Figure 19. Extended legend in next page

Figure 19. **Niche characterization and pathway analysis in MS lesions.** Repurposed from Lerma-Martin *et al*. 2024. **a**, Cumulative number of DEGs per niche between MS lesion types. Color indicates MS lesion type. **b**, Heatmap of pathway enrichment activities between niches. Color indicates association to a MS lesion type. **c**, Volcano plot of DEGs for PPWM, LR, LC and VI niches. Color indicates association to a MS lesion type. **d**, AS and MC compositional changes per individual niches between MS lesion types. Asterisks indicate significance (two-tailed Wilcoxon rank-sum test, BH adjusted P < 0,10).

Notably, the niches enriched for TGFβ signaling also showed a loss of SMAD2/SMAD3 function, the direct transcription factors of this pathway, suggesting a potential signaling cascade dysfunction, particularly in the LC and LR niches (**Fig. 19b**). These pathway activity findings provide insight into how different cell communities are responding to disease conditions.

Next, I proceeded to then examine gene changes within these niches, as they could enhance the understanding of specific functions occurring in MS-CA versus MS-CI lesions. In the previous analysis, I observed that the LR and LC niches in MS-CA exhibited the most significant gene changes. Within the LC niche, there was presence of genes that have been identified to contribute to activation of MC cells into a neuroprotective phenotype (*MAFB, VISG4*), as well as genes that promote tissue homeostasis and repair (*MERTK, PHGDH*). In contrast, the LR niche, while also containing genes associated with anti-inflammatory functions (*TGFBR2*), shared several genes with the VI niche. These shared genes, involved in cell migration and invasion (*FSCN1, LAP3*), proliferation (*PAG1, S100A9*), and inflammation (*SPP1, C1QA, HSPH1*), underscoring the interrelated functions of LR and VI niches. Additionally, the VI niche exhibited a unique gene expression profile, including genes associated with cell differentiation (e.g., *MOB1A*) and epithelial-to-mesenchymal transition (e.g., *CCT5*). These distinct functions of these genes further highlight the VI niche as an important target for deeper investigation (**Fig. 19c**).

In the PPWM, I observed gene expression patterns related to inflammation (e.g., *S100A8*) and lipid regulation (e.g., *UGCG*). However, these patterns shift in MS-CI lesions, where the PPWM is characterized by genes associated with cytoskeleton remodeling and functions such as cell motility, inhibition of proliferation, and cell survival (e.g., *ITGA7, AHNAK, FAT1, ST5, CDC42EP4, PLEC*) (**Fig. 19c**).

As the final analysis of this section, with the collaboration of Pau Badia i Mompel, we quantified the changes in abundance of niches and associated cell types in space (**Methods**). While the overall abundance of specific cells in niches did not significantly change between MS lesion types, AS were notably enriched in MS-CI relative to MS-CA lesions, suggesting the build-up for the astroglial scar in the demyelinated core. Moreover, MC cells demonstrated enrichment in the LR

of MS-CA, suggesting a sustained inflammation at the rim (**Fig.19d**). Taken together, these findings provide a detailed characterization of the distinct niches present in both lesion types, offering an overview of lesion progression at a high spatial resolution.

4.4 Mapping of cell-cell interactions in MS

4.4.1 Computational pipeline analysis for cell-cell interactions

In this final chapter, I aim to delve into the complex network of interactions occurring within the inflammatory milieu of MS lesions. This environment is shaped by the interplay of various cell types, and understanding them is key to comprehending the dynamics of lesion progression. To achieve this objective, I will use all insights that gained through this study, and together with Pau Badia i Mompel, we will devise a novel approach to explore cell-cell communication (CCC) events in the form of ligand-receptor interactions.

Our initial focus was addressing the high number of false positive results coming from CCC inference solely based on single cell data[187]. To overcome this issue, we devised a computational approach that would infer CCC events based on our paired snRNA-seq atlas and ST data. Spatial information has been shown to significantly improve results by increasing the likelihood of possible CCC through colocalization events, where the physical proximity of cells enhances the probability of true interactions and acts as a filter of possible false positives[188].

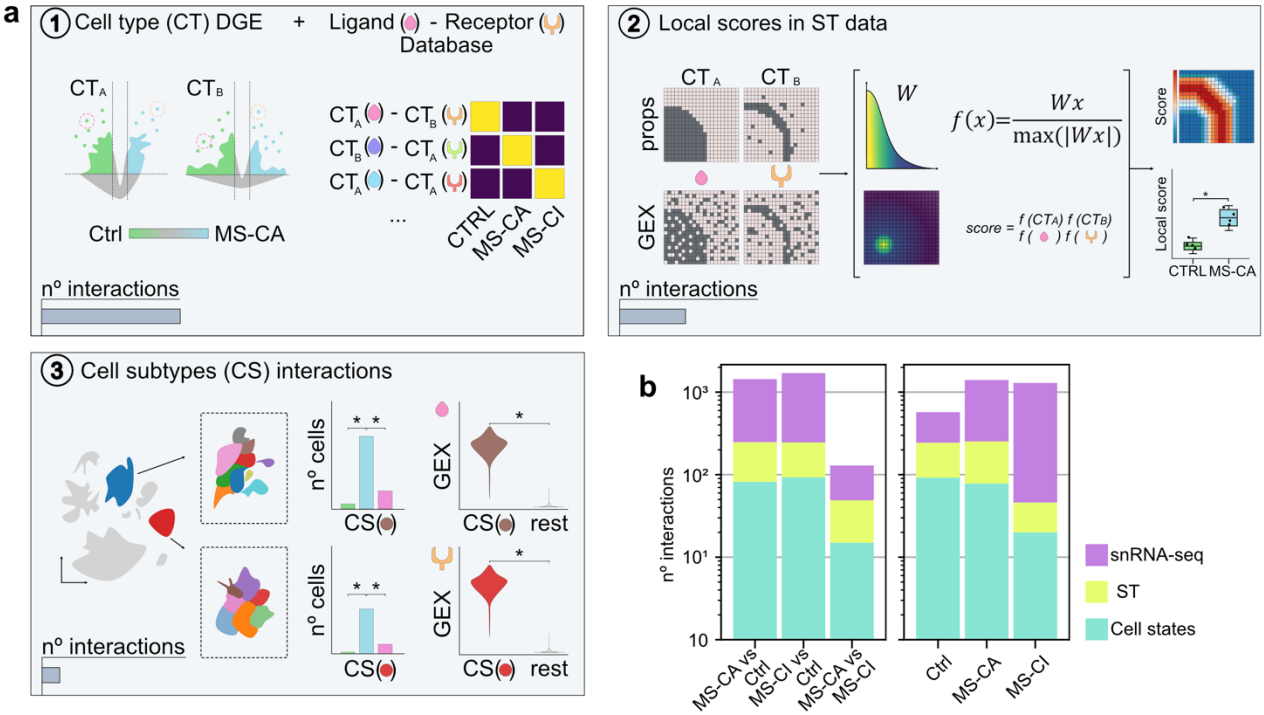


Figure 20. Extended legend in next page

Figure 20. Summary of cell-cell communication pipeline workflow. Repurposed from Lerma-Martin *et al*. 2024. **a**, Pipeline design for the inference and filtering of cell-cell communication (CCC) events. In the first step (1), differential expression analysis results obtained from snRNA-seq were used to obtain significant cell type pair ligand-receptor interactions (BH adjusted P < 0,15). In the second step (2), spatially weighted interaction local scores are computed for the selected interactions and tested for significance across lesion types (BH adjusted P < 0,15). In the third step (3), interactions are only kept if their ligand or receptor was a marker gene for a MS lesion type and specific cell subtype. Note at each step, conflicting interactions were dropped off to ensure robustness. **b**, Number of cell-cell interactions obtained at each step of the pipeline.

The approach consisted of a three-step process, where we progressively tightened our filtering criteria. First, leveraging the differentially expressed genes identified in Chapter 3, we calculated significant ligand-receptor interactions among all different cell types pairs within each MS lesion. Next, we computed a spatially weighted local score per spot in the ST data based on these interactions, selecting only those that remained significant across conditions. Finally, we kept the interactions only if either the ligand, the receptor or both were specific to a cell subtype significantly abundant in the same lesion type (**Fig. 20a, Methods**). Through the execution of this pipeline, we successfully predicted 190 differential CCC interactions from the original pool of 2.724 (**Fig. 20b, Supplemental Table 3-5**).

4.4.2 Interaction patterns and spatial context

Among these 190 interactions, I examined how many were unique to each condition and the specific cell types they were enriched in. Notably, there was no overlap between CTRL and MS lesions, indicating distinct interaction trends (**Methods**). Overlap between MS-CA and MS-CI lesions was low, indicating that the CCC events were condition specific (**Fig. 21a**).

Upon closer examination at the cell types with the highest number of differential interactions, OL, AS and MC were again ranked at the top three (**Fig. 21b**). However, the nature of their cell type-specific crosstalk varied between CTRL and MS (**Fig. 21**). In CTRL samples, interactions were predominantly associated with OL, though the network also included OPC, AS, MC, and EC, reflecting a heterogeneous pattern (**Fig. 21d**). In contrast, in MS lesions AS and MC were central to the interactions, with a network predominantly driven by immune cell types interacting with AS, SC, and EC (**Fig. 21e-f, Methods**).

With a preliminary understanding of how the cell types were interacting, I next sought to identify the specific interactions taking place. Many interactions were not straightforward, involving multiple scenarios where, for example, a cell expressed the same ligand that interacted with different cells having the same receptor.

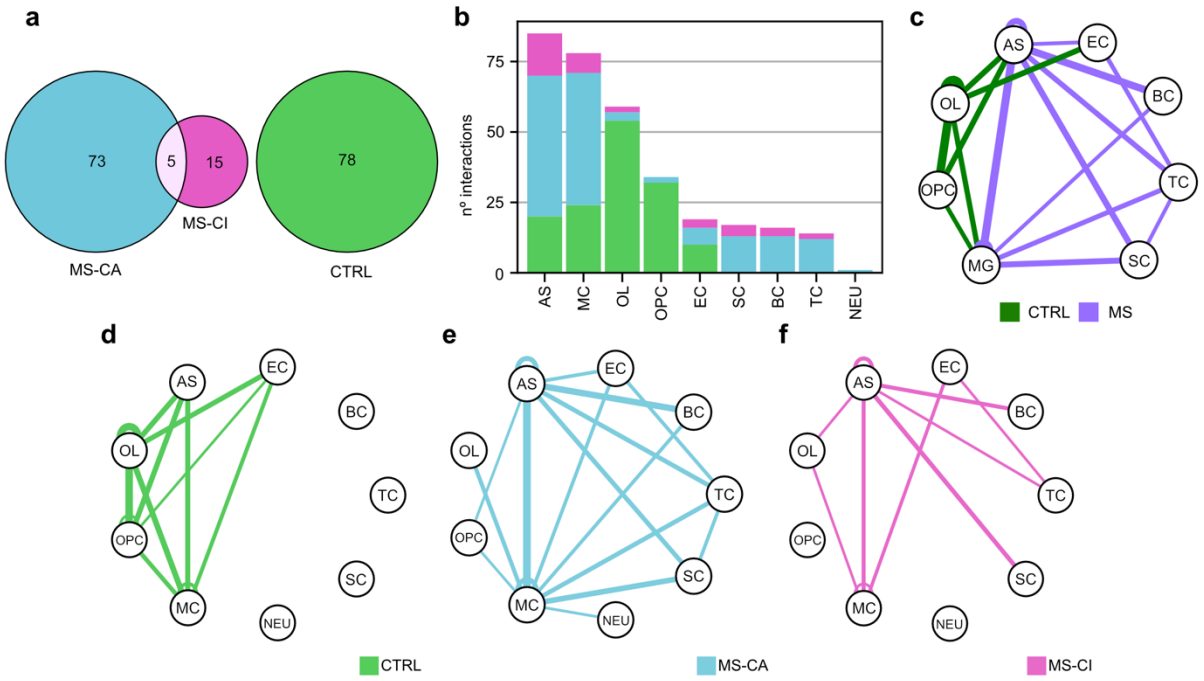


Figure 21. **Overview of interactions between conditions and cells.** Repurposed from Lerma-Martin *et al*. 2024. **a**, Venn diagram of overlapping interactions between MS lesion types and CTRL. **b**, Cumulative number of differential interactions grouped by cell type across conditions. Color indicates condition. **c,** Cell-cell network structure. Edge size indicates number of interactions. Green indicates CTRL and purple both MS conditions. **d**, **e**, **f**, Individual cell-cell network structures. Edge size indicates number of interactions. Color indicates condition.

Visualizing these complex networks can be challenging due to the numerous factors involved. To simplify this, Pau Badia i Mompel and I designed a heatmap that summarized these interactions numerically and indicated the specific conditions in which they occurred (**Fig. 22a**). For example, the interaction between PSEN1 and NOTCH1 is represented as PSEN1 ligand expressed by OL interacting with three different receptor cells expressing NOTCH1: AS, MC, and OL. This is indicated by the numbers 3 - 1:1:1, indicating one interaction with each of the three cell types.

To further understand the spatial context of these interactions, we mapped them to their respective niches within the tissue. By incorporating spatially restricted local scores, we were able to predict where these interactions were most likely to occur. In MS-CA, the majority were mapped to the LR niche, while in MS-CI, interactions were associated with both the VI and LR niches (**Fig. 22b**). To ensure the accuracy of these predictions at both the gene and spatial levels, I validated five interactions predicted to take place at the inflamed rim of MS lesions, a critical niche in disease.

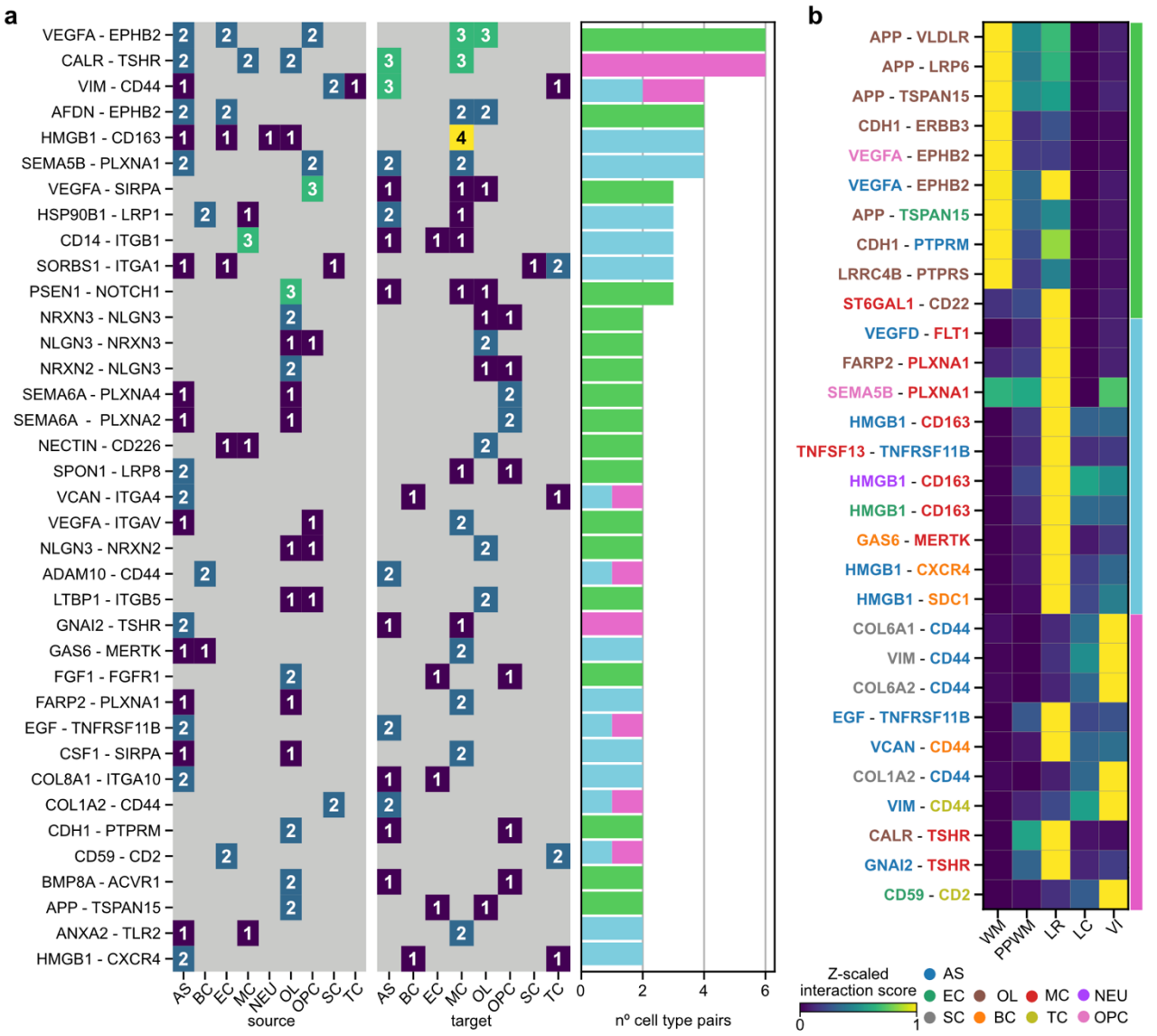


Figure 22. **Visual summary of multicellular interactions**. Repurposed from Lerma-Martin *et al*. 2024. **a**, Heatmap showing multicellular ligand-receptor interactions. Numbers indicate the number of cell type pairs. Color indicates association to MS lesion type. **b**, Top 30 significant interactions per MS lesion types and CTRL across niches. Text color indicates sender (left) and receiver (right) cell types. Column on the right represents condition.

4.4.3 Spatial mapping of HMGB1 interactions with CD163 and TLR2

High mobility group box 1 (HMGB1) is a nuclear protein that, upon infection or injury, can be secreted into the extracellular space, acting as a prototypical damage-associated molecular pattern (DAMP) molecule. It plays a crucial role in regulating diverse inflammation and immune responses, depending on the receptor that it interacts with[189]. Studies have shown that when *HMGB1* interacts with the MC-specific receptor *CD163*, it elicits an anti-inflammatory

response[190,191]. However, interactions with members of the toll-like receptor (TLR) family, such as *TLR2*, may activate an proinflammatory response by stimulating the innate immune system[189].

In the data, the HMGB1 ligand was predicted to participate in CCC events across several cell types, including AS, NEU, EC, and OL. Its corresponding receptors, CD163 and TLR2, were expressed in MC cells. Computational analysis revealed a significant upregulation of *HMGB1-CD163/TLR2* interaction between AS and MC in the LR of MS-CA samples (**Fig. 23a-c, 24a-c**). With the enhanced resolution provided by our pipeline, we predicted these interactions to occur between the AS subtype Dis1 and the three MC subtypes CA, Dis and Rim. To validate these findings, I performed smFISH, using probes targeting *ADCY2* to identify AS, *CD163* or *TLR2* to identify MC, and *HMGB1* to identify the ligand (**Methods**). I was able to observe double positive *ADCY2*/*HMGB1*-expressing AS cells in close spatial proximity to *CD163*- or *TLR2*- expressing MC cells (**Fig. 16c, f**). Notably, these events were highly specific for LR areas of MS-CA lesions, thus strengthening the validity of our predictions.

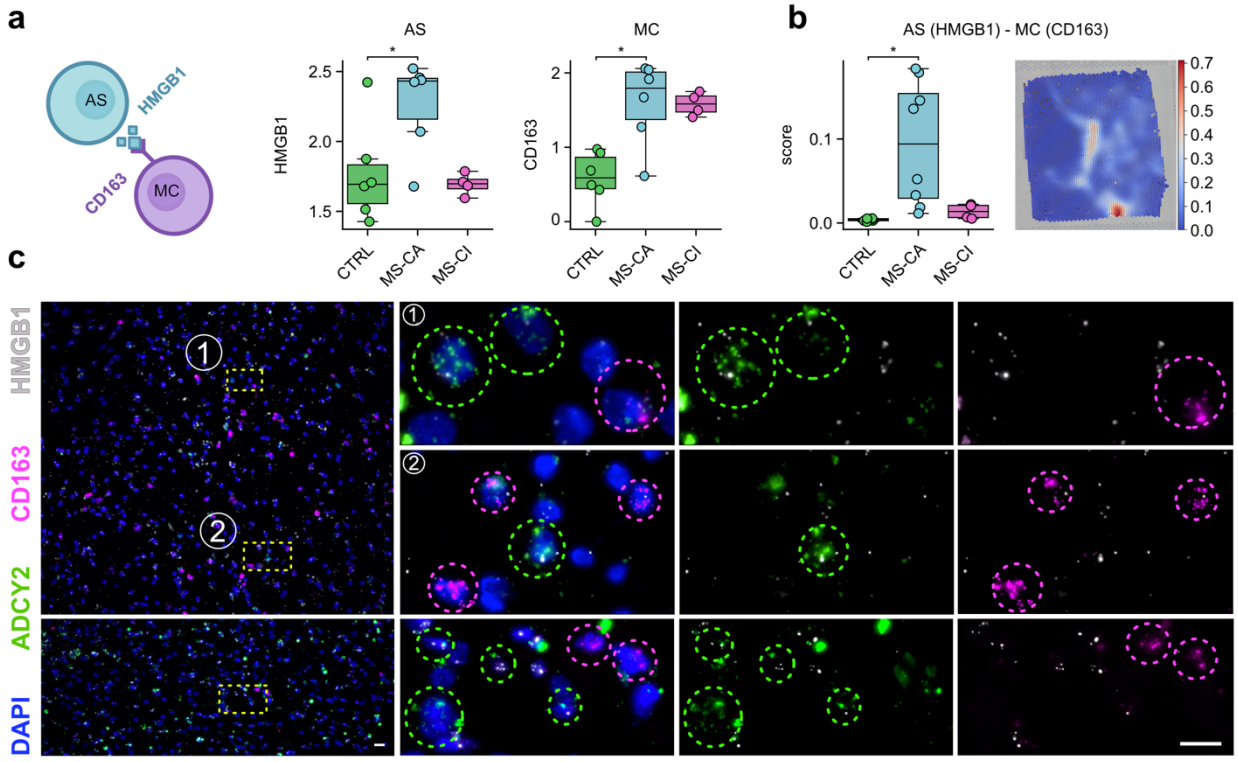


Figure 23. Validation of predicted HMGB1-CD163 interaction at the LR of MS-CA lesions. Repurposed from Lerma-Martin *et al*. 2024. **a,** Boxplot of HMGB1 (ligand) and CD163 (receptor) in AS and MC, respectively, between conditions. **b**, Boxplot showing cell-cell interactions scores between ligand-receptor together with predicted ST mapping of the interaction. **c**, smFISH for *ADCY2* (AS), *CD163* (MC) and *HMGB1* (ligand). **d-f**, Same plot distribution for HMGB1 (ligand) and CD163 (receptor). Asterisks indicate significance (Wald test, BH adjusted P < 0,05). Scale bars: 20µm.

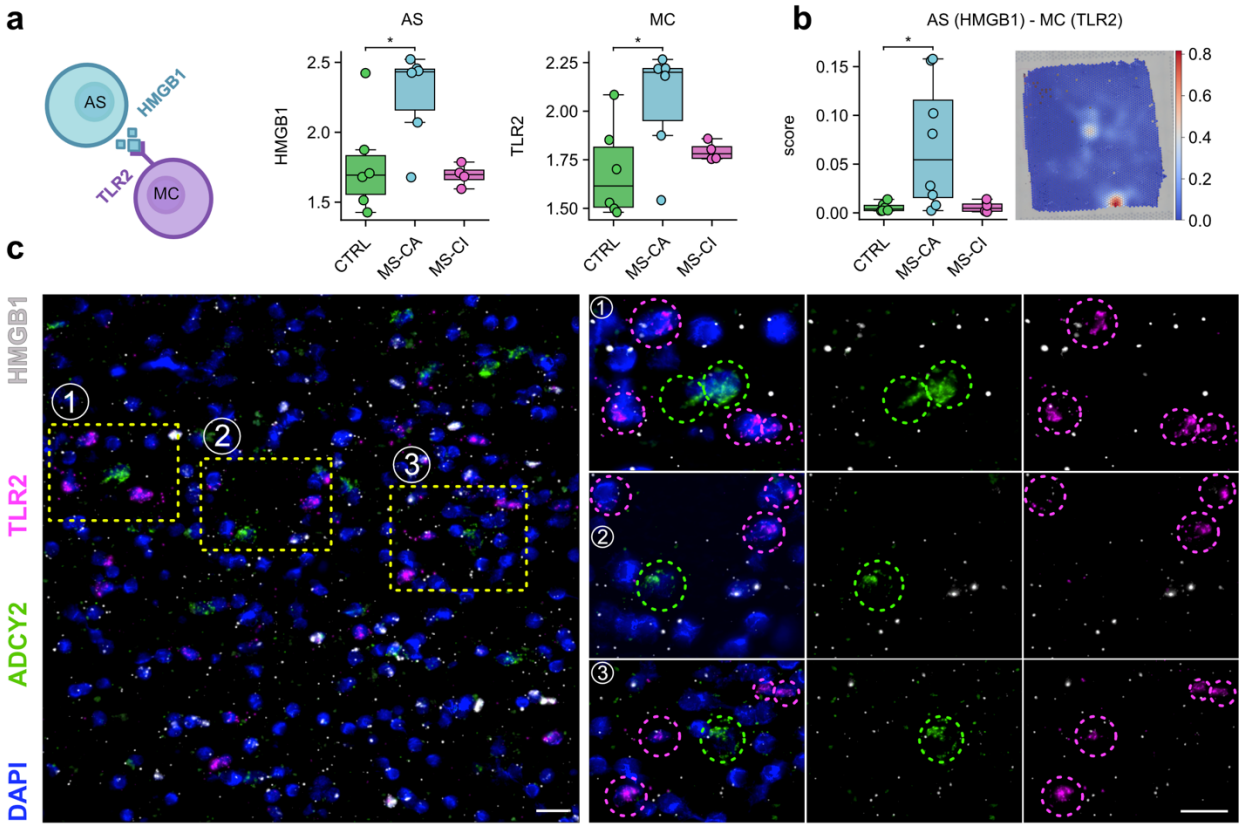


Figure 24. Validation of predicted HMGB1-TLR2 interaction at the LR of MS-CA lesions. Repurposed from Lerma-Martin *et al*. 2024. **a,** Boxplot of HMGB1 (ligand) and TLR2 (receptor) in AS and MC, respectively, between conditions. **b**, Boxplot showing cell-cell interactions scores between ligand-receptor together with predicted ST mapping of the interaction. **c**, smFISH for *ADCY2* (AS), *TLR2* (MC) and *HMGB1* (ligand). **d-f**, Same plot distribution for HMGB1 (ligand) and TLR2 (receptor). Asterisks indicate significance (Wald test, BH adjusted P < 0,05). Scale bars: 20µm.

4.4.4 Spatial mapping of CD14 and ITGB1 interaction

CD14 is a gene that encodes for a receptor specifically expressed by MC cells. Its main function is to mediate the innate immune response upon binding to a bacterial-derived lipopolysaccharides (LPS). Additionally, CD14 can also act as a surface ligand, although its functions in this capacity have been less extensively documented[192]. *ITGB1*, also known as *CD29*, encodes for an integrin that acts as a cell surface receptor. Depending on the ligand it interacts with, ITGB1 has been associated with different biological functions like development, cell homeostasis or even innate immune responses. However, little is known about the function associated with its binding to CD14[192].

Our pipeline predicted a *CD14-ITGB1* CCC event between MC, AS and EC, which was spatially mapped to the LR niche of MS-CA samples. As before, we could trace the interaction down to the

distinct cell subtypes predicted to be responsible for them. *CD14* was expressed as a ligand in MC Dis and Rim, while *ITGB1* was expressed as a receptor in MC Dis, AS Dis1, AS React and EC Prolif. To validate these interactions, I performed smFISH, using probes targeting *CD14* to identify MC, *ADCY2* to identify AS, *VWF* to identify EC and *ITGB1* to identify the receptor (**Methods**). Of note, I observed MC, AS and EC cells expressing *ITGB1* in close proximity to MC cells expressing *CD14*, validating a possible interaction between a surface ligand and its receptor (**Fig. 25a-c, 26, 27**). In agreement with the prediction, I did not observe those interactions in CTRL samples based on smFISH validation.

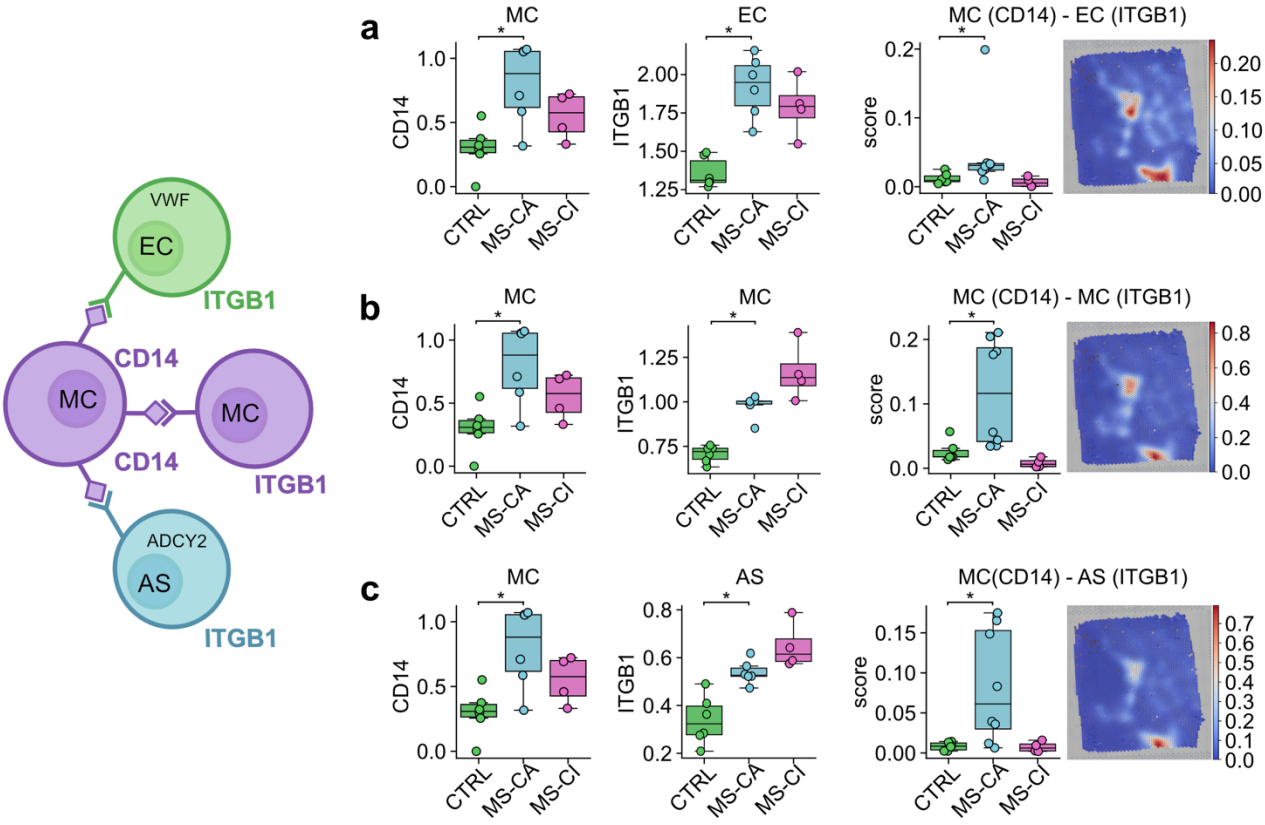


Figure 25. **Spatial mapping of predicted CD14-ITGB1 interaction.** Repurposed from Lerma-Martin *et al*. 2024. **a,** Boxplot of CD14 (ligand) from MC and ITGB (receptor) from EC (top row), MC (middle row) and AS (low row) between conditions. Cell-cell interactions scores between ligand-receptor together with predicted ST mapping of the interaction (right in all panels).

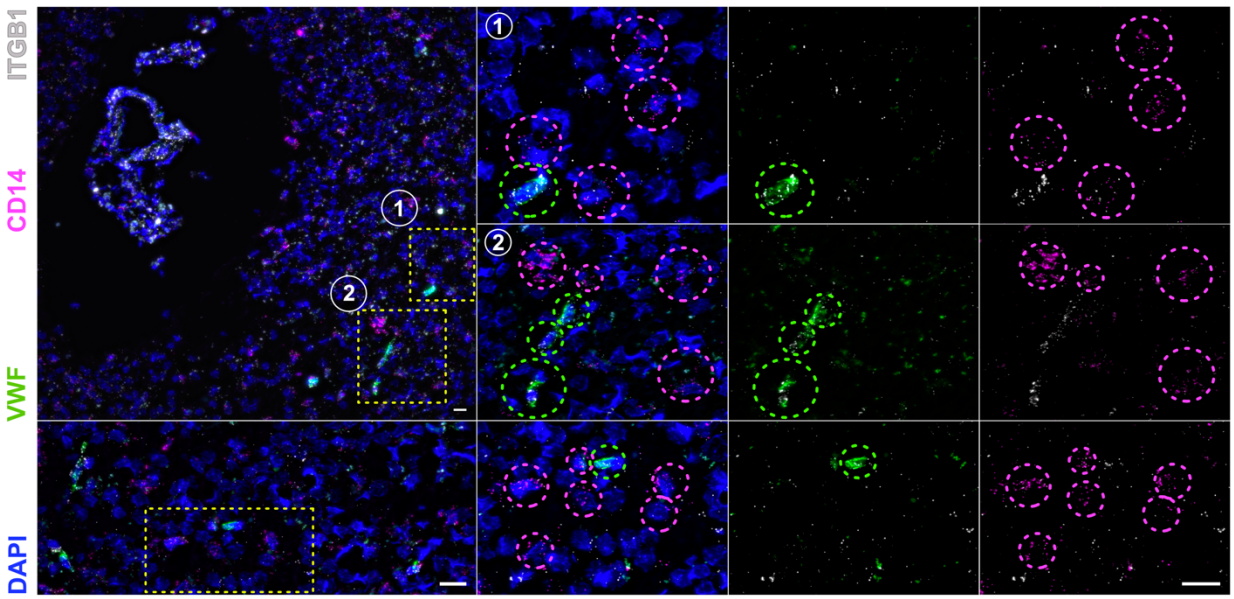


Figure 26. **Validation of predicted CD14-ITGB1 interaction in EC and MC cells at the LC of MS-CA lesions.** Repurposed from Lerma-Martin *et al*. 2024**.** smFISH for *CD14* (MC), *ITGB1* (EC/MC) and *VWF* (EC).

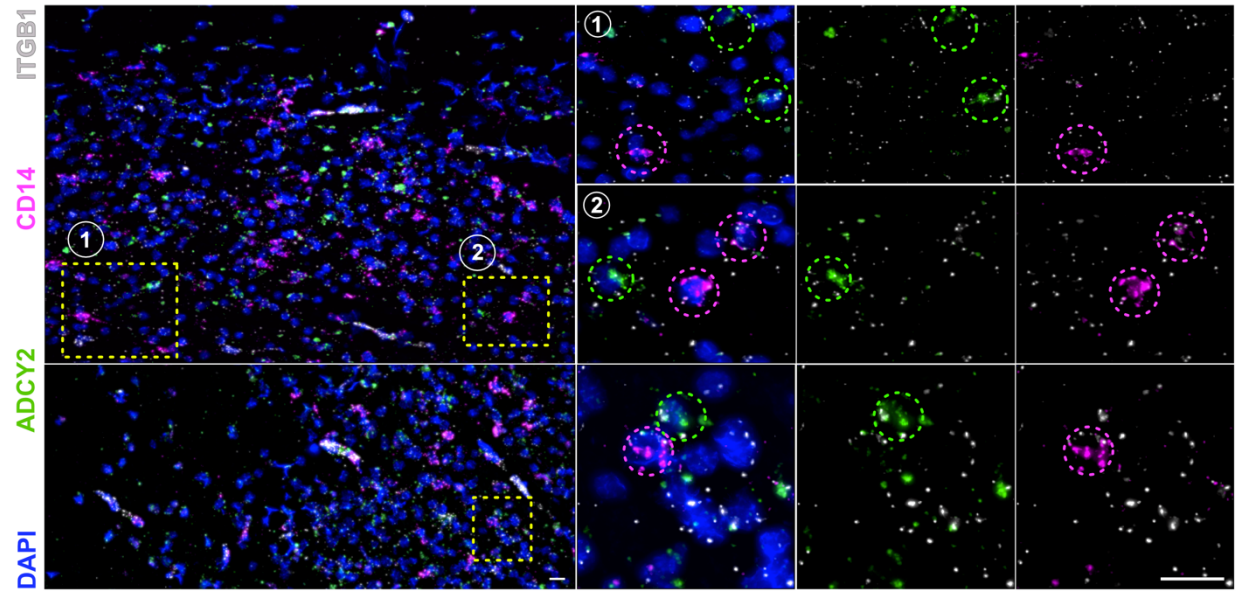


Figure 27. Validation of predicted CD14-ITGB1 interaction in AS and MC cells at the LC of MS-CA lesions. Repurposed from Lerma-Martin *et al*. 2024**.** smFISH for *CD14* (MC), *ITGB1* (AS/MC) and *ADCY2* (AS). Scale bars: 20µm.

In summary, in this chapter I introduced an unsupervised approach for predicting potential cell-cell communication events associated with restricted niches in multiple sclerosis lesions, that I subsequently validated. This approach could potentially be used as a tool for investigating new MS stage-specific biomarkers or potential therapeutic targets.

5. Discussion

Multiple sclerosis (MS) remains a complex, multifaceted disease with many unresolved questions, ranging from its etiology to how lesions progress. Until recently, MS pathology research primarily relied on conventional histopathological techniques, which, though valuable for answering specific questions, were limited in scope[1,9,10]. These methods require numerous antibodies to distinguish cell types, which limits their flexibility to explore broader aspects of the disease. More recently, imaging-based spatial transcriptomics, such as *in situ* hybridization, introduced new layers of resolution[193] but still rely on a set number of probes, making it challenging to fully capture the complexity of the resulting lesions.

Technological advancements over the past decade have revolutionized our ability to explore MS in greater detail. The introduction of single-cell and nucleus RNA sequencing has provided unprecedented resolution, allowing for a deeper understanding of cell-type-specific gene expression and its alterations[113,114,194–196]. By integrating this cutting-edge technique with spatial transcriptomics, we can now examine gene expression in specific cell types alongside their spatial contexts. This approach can be further supported by traditional *in situ* imaging, which helps validate the results and provides a complete view of tissue structure. In this study, I used this integrated approach to create a high-resolution transcriptomic dataset of brain white matter in both control and MS samples[159]. This dataset allowed me to decipher intricate tissue niches and cell-cell interactions at specific locations within the tissue. Such comprehensive analyses provide a deeper understanding of MS pathology and the cellular mechanisms driving disease progression.

My initial focus was on creating a comprehensive atlas, as it would serve as the main building block for my subsequent analyses. Due to its critical role, the atlas was highly curated and annotated, and I ultimately identified 9 main cell types and a total of 79 unique cell subtypes, each with its distinct phenotype. This in-depth characterization is one of the first main results of this study, as one of my main goals is for this data to be used by the community. Existing atlases often suffer from insufficient annotation, only highlighting the cell types relevant for their specific studies, resulting in many interesting cell types remaining unexplored.

Furthermore, I performed spatial transcriptomics on 18 samples, of which 15 were paired with the samples used in the snRNA-seq atlas. I chose to use the Visium platform, which does not provide single-cell resolution, so we inferred cell type composition per spot using the atlas as reference.

This combined approach allowed me to map all main cell types and their subtypes across different slides, providing detailed insights into tissue architecture in MS lesions. Of note, the observed patterns aligned with the biological expectations, validating our findings and complementing recent probe-based *in situ* sequencing studies[193].

Currently, there is a paucity in publicly available datasets from cross-condition setups[197–199], resulting in many unexplored methods for analyzing these extensive datasets. I aimed to develop a novel approach for analyzing paired data by the community while also enhancing our understanding of lesion progression in multiple sclerosis. As MS lesions become chronic, cells assume different key roles in response to various activation stimuli, forming distinct cell communities and leading to compartmentalized[164] inflammation and tissue damage. Detecting these communities and understanding their cellular composition and spatial relationships can provide valuable insights previously unattainable.

In this project, I collaborated closely with neuropathologists to obtain annotations of the different tissue areas based on existing knowledge. I used these annotations to validate the unsupervised factor analysis that Pau Badia i Mompel and I developed for characterizing distinct niches. This method integrates multiple data views to identify unique variability and create factors for characterizing distinct cell communities or niches in both control and disease states. I identified four niches unique to disease and present in both lesion types: vascular infiltrating (VI), lesion core (LC), lesion rim (LR), and periplaque white matter (PPWM).

This unsupervised approach identified a biologically significant niche, the VI niche, which was not detected by the manual annotation. It primarily consisted of T-cells (TC), stromal cells (SC), and endothelial cells (EC), with minor contributions from B-cells (BC) and myeloid cells (MC). This niche is of particular interest as it can help understand one of the potential mechanisms behind the formation of white matter lesions. ECs, a key component of the VI niche, are crucial for maintaining the integrity of the blood-brain barrier (BBB). These cells are tightly connected by interendothelial junctions, regulating molecular and cellular trafficking between the blood and the central nervous system (CNS). Damage or stress on these cell-cell contacts leads to BBB dysfunction, resulting in hyperpermeability and the infiltration of immune cells into the perivascular spaces of the brain[200].

In addition to ECs, stromal cells (SC) contribute to the structural and functional support of the BBB and regulate immune responses within the CNS[201]. In the context of multiple sclerosis (MS),

when autoreactive lymphocytes primed against CNS-reactive antigens in the periphery enter the CNS, they become reactivated in these perivascular spaces[202]. At these sites, SC cells produce chemokines that attract and retain leukocytes, leading to their accumulation[203,204]. These cells not only support the proliferation of immune cells but also enhance the production of both pro- and anti-inflammatory molecules[205]. Additionally, stromal cells elaborate a fibroblastic reticular network that interacts with the infiltrating lymphocytes at white matter lesions[206], further contributing to the inflammatory environment.

The spatial mapping of the VI niche suggests that smoldering tissue inflammation may result from BBB leakage, a conclusion supported by the gene enrichment analysis of MS-CA lesions, which shows gene functions related to cell migration, invasion, proliferation, and inflammation. This analysis not only supports previously suspected biological mechanisms but also provides valuable insights into the cellular components, phenotypes, and pathway activations within this critical tissue niche in MS pathology.

The VI niche was predominantly located within the lesion core (LC) niche, which is enriched with astrocytes, essential for the formation of the astroglial scar in MS lesions. The gene functions enriched within the LC niche, which are predominantly related to anti-inflammation, tissue regulatory functions, and remyelination support, provide further evidence of active scarring processes.

During the characterization of the LC niche, I identified a previously unreported astrocyte subtype, the white matter motile-like ciliated astrocytes. Although astrocytes with primary or non-motile cilia have been identified in most CNS regions containing gray matter in animal models, their length and abundance varies[207–212]. The main function of these cilia is to transduce molecular signals from the environment[211,213,214], and they are believed to regulate numerous neuronal functions, although their precise role is still unknown[208].

Additionally, I found the ciliated astrocytes mapping to the lesion core of subcortical MS lesions, expressing genes characteristic of motile cilia (FOXJ1, DNAH11, SPAG17, CFAP54)[183,184,186,215], and with cilia lengths ten times longer than previously reported. Although I have not yet performed any functional characterization beyond pathway enrichment, I hypothesize that this cell subtype, which lacks proinflammatory profile, may contribute to astroglial scar formation. Interestingly, primary cilia have been reported in wound repair and scar formation in various tissues[216–219], but not the motile type. Given the densely packed structure of the brain, it is unlikely that these cilia are

functionally motile. This raises the possibility that genes traditionally associated with motile cilia might also play a role in non-motile structures under certain microenvironmental disruptions. One could hypothesize that, as seen in other studies[220], these motile cilia genes might help regulate the cilia waveform by contributing in the formation of a structure that enables the cilia to extend to lengths previously out of reach by non-motile cilia. Once they cover enough area, these cilia could act as hubs for propagating microenvironmental cues and function as signaling modules. If this is the case, this structure would be crucial for astrocytes to communicate with the other cell types under stress conditions and build the glial scar. This opens up a line of study in the MS research field and other diseases or injuries that lead to scar formation in CNS tissues, such as strokes, where recent evidence suggests the presence of this cell subtype.[221]

Undeniably, tissue niches play a pivotal role in shaping the pathophysiology of MS lesions, influencing cellular responses and adaptations, helping us understand better the immune cell infiltration events, and ultimately shedding a light disease progression. For tracking lesion evolution, it has been suggested that the presence of a lesion rim (LR) in MS-CA lesions serves as a key indicator due to its high inflammatory activity[222]. By characterizing the LR niche, mainly enriched with myeloid cells (MC), I observed that enriched gene functions were related to cell migration and invasion, proliferation, inflammation, antiviral responses and tissue remodeling. Interestingly, this LR niche was also identified in MS-CI lesion, although it was not annotated by the neuropathologists. MS-CI lesions are characterized by a lower number of activated MC, indicating reduced inflammatory activity, and a halt in ongoing tissue damage[1]. This observation aligns with my results, which show that the enriched pathways are related to tissue remodeling and repair. By examining the gene expression patterns within these niches and the cells that constitute them, both myself and other researchers that use this data can gain valuable insights into the molecular mechanisms underlying lesion development and progression.

The study of temporal dynamics of lesions with human tissue poses a challenge as I could only capture a snapshot of the disease at a single point in time, limiting my ability to track how lesions evolve. To address this, I correlated the gene expression changes to specific cell subtypes to understand how lesions evolve. All cells exhibited high plasticity, adapting to both physiological and pathological conditions. This adaptability was detected through our approach, revealing that some cells displayed traits typically associated with earlier, less differentiated forms, suggesting a potential reversal to a stem cell-like state. An interesting example would be the ECs, which in diseased tissue express embryonic-like genes, such as those involved in endothelial-to-hematopoietic transition (*MEIS2*)[223] and endothelial-to-mesenchymal transition (*CCT5*)[224],

processes crucial in development and wound healing[225]. Similarly, astrocytes (AS) displayed this regressing phenotype with ciliated astrocytes, which have been observed during early gestational weeks in human fetal spinal cords but later disappeared during development[180]. In addition, some studies have characterized a phenotype in AS after brain injury that resembles the Reactive AS in my data. These AS express a crucial gene (*LGALS3BP*), which is associated with acquiring neural stem cell-like properties following injury and when exposed to blood[226], which could be consistent with the BBB leakage scenario. This suggests that certain stressors in the inflammatory environment may induce phenotypic plasticity, allowing some cells to revert to earlier developmental stages, while others adapt without fully regressing. Exploring the specific triggers that cause this reversion and understanding why other cells may not transition into a stem-like state, or if they could under different conditions, could provide valuable insights.

To further elucidate the mechanisms underlying lesion progression, I explored the cell-cell interactions occurring within the different niches in both CTRL and disease conditions, as these are crucial for interpreting phenotypic behaviors and adaptations observed in various cell subtypes. To achieve this, Pau Badia i Mompel and I developed a computational method that integrated both data modalities, focusing on the ligand-receptor interactions that were significant in snRNA-seq and spatially proximal in the ST data. Due to our stringent filtering criteria, we predicted a small number of interactions, which we were also able to map back to their respective niches. About half of these interactions were present in control conditions but absent in disease, or specific to MS-CA lesions. I validated some interactions that had been previously characterized in the context of inflammation, along with others that had not been reported in the brain before, demonstrating the novelty and robustness of the findings.

As a well-documented inflammatory event, I focused on HMGB1, a central inflammatory mediator produced by AS, OL, NEU and EC in the dataset. I validated interactions between AS-derived HMGB1 and two receptors on MC cells. The first receptor, CD163, is a haptoglobin scavenger known for its anti-inflammatory roles[227], selectively upregulated in MC cells at the rims of MS-CA lesions, where it serves as a key mediator for iron uptake[190]. Additionally, CD163 has been identified in other studies as an anti-inflammatory receptor for HMGB1-haptoglobin complexes[191], reinforcing its role as a moderator of inflammation in MS-CA lesions. This finding highlights the role of CD163-positive MC cells as regulatory MC, as they not only facilitate iron uptake but also mitigate inflammation by neutralizing damage-associated molecules. In contrast, the second receptor that I validated, TLR2, has been suggested as a key initiator of neuroinflammation upon interaction with HMGB1[228]. Recent studies suggest that this interaction may impair OPC

proliferation following injury, preventing their maturation into OLs and affecting axonal remyelination, thereby presenting a challenge to the repair process[229]. These interactions shed light on the complex dynamics within MS lesions, revealing how different receptors from the same cell when interacting with the same ligand can either moderate or exacerbate inflammation.

Another interaction, previously reported only in colorectal cancer and associated with chemotaxis, involves the MC-expressed CD14 ligand and ITGB1 receptor[230]. This interaction is not well characterized in neuroinflammation, nor has the interaction between MC-expressed *CD14* and AS- or EC-expressed *ITGB1* been thoroughly described. As a result, the implications of these interactions for disease remain speculative. Previous studies suggest that AS progenitor cells use ITGB1 as support for blood-vessel-guidance migration[231], potentially enabling AS cells to move closer to inflamed areas in MS lesions establishing the glial scar. This function aligns with their observed regressing phenotype and conversion to stem-cell-like properties under tissue injury, reflecting their adaptability during disease progression. Moreover, ITGB1 in EC cells is linked to both the stabilization[232] and disruption of blood vessel integrity,[233] depending on the context, making it a promising target for understanding immune cell infiltration. Although the exact role of these interactions is still uncertain, it introduces a new dimension to the understanding of lesion dynamics.

In addition to these validated interactions, other predicted interactions could offer insights, especially regarding immune-vascular interactions between EC and infiltrating lymphocytes and immune cells. Several of these involve EC-expressed ligands that may bind to integrins related to cell adhesion and migration, and receptors in lymphocytes essential for immune surveillance and migration to inflammation sites. One such interaction is between TGM2[234] and VWF[235] from EC with ITGA4 from BC[236] and TC[237,238] cells. Another relevant example involves EC-derived HLA-E interacting with the T-cell receptor CD8A. The ability of HLA-E to modulate CD8 T cells in a context-dependent manner, either enhancing or limiting their cytotoxic activity[239], adds complexity to the understanding of cell infiltration. These findings, along with other validated examples and predicted cell-cell communication events, demonstrate the utility of these computational methods in identifying biologically relevant interactions. They also highlight how these interactions can be linked to specific cell subtypes and tissue niches in MS.

6. Future prospects and conclusions

In this study, I aimed for a well-balanced cohort in terms of sex, age, and batch. However, the limited availability of high-quality tissue restricted my sample size for transcriptomic analyses or techniques like *in situ* hybridization. While new techniques are emerging, they currently lack the resolution of other established methods. Future advancements that enable high-quality analyses from lower-quality RNA samples would greatly benefit the research community.

The scarcity of suitable tissue limited my ability to study acute lesions and compare multiple lesions within individual patients. These research areas could offer valuable insight into lesion development and identify factors which are patient-specific or inherent to the disease itself. The acquisition of high-quality tissue samples in the future will enable us to pursue these promising research lines further.

Overall, in this study, I have generated a substantial amount of data that is now available for further exploration by the research community[159]. While snRNA-seq has established standards and workflows for quality control and analysis, ST remains an emerging field with less consensus on advanced analytical methods beyond deconvolution. By sharing the approach and insights that Pau Badia i Mompel and I have developed, I aim to advance the field and encourage further exploration and utilization not only of ST data but also of integrated paired atlases.

This approach, which integrates computational prediction models with *in situ* validation at the RNA and protein levels, provides a robust framework for studying MS pathology. However, functional validation and further investigation will be crucial to fully understand and expand upon these findings, such as the several newly identified cell-cell communication events in the brain and the discovery of ciliated astrocytes. Characterizing these cells and performing functional studies will help reveal their role in disease and the triggers for cilia development, paving the way for new research avenues

In summary, I have successfully generated and computationally analyzed a paired single-nucleus and spatial transcriptomics dataset, enabling us to better understand the intricate tissue microenvironment involved in MS lesion progression. Through this approach, I identified biologically relevant tissue niches, novel cell subtypes, and unique cell-cell interactions occurring within specific regions of the lesions. With this work, I aim to contribute to a deeper understanding of this complex disease and pave the way for the discovery of potential therapeutic targets.

7. Acknowledgements

First of all, I want to express my deepest gratitude to everyone who played a crucial role and stood by my side throughout this journey. Although it has been challenging at times, it has also been incredibly rewarding, and I will always be grateful for your unwavering support.

Prof. Dr. Lucas Schirmer, I would like to thank you for the opportunity to pursue my PhD within your group. Your trust in allowing me to implement the first spatial techniques in the lab gave me the chance to truly make them my own. I am very thankful for your scientific support throughout my PhD journey, as well as for your belief in my ideas, which allowed me the freedom to explore and develop them independently. Your encouragement to attend courses and international conferences helped me gain a deeper understanding of the field and gave me the confidence to present my project in front of large audiences. Although my journey was challenging at times, I am grateful for your guidance, as it has allowed me to achieve my goal of becoming a doctor. For all of this, I am extremely thankful to have had you as my supervisor.

To all the members of my thesis advisory committee: **Prof. Dr. Ana Martin-Villalba, Prof. Dr. Lucas Schirmer** and **Prof. Julio Saez-Rodriguez**. Thank you for your time, support, and valuable feedback. Your insights and comments during our TAC meetings were essential in keeping me on track throughout my PhD journey.

To all of the members of my PhD defense committee, thank you for taking the time to participate in my defense and grade my work. **Prof. Dr. Lucas Schirmer**, thank you for agreeing to be my examiner and proofreading my PhD thesis beforehand. **Prof. Dr. Ana Martin-Villalba**, thank you for chairing my PhD defense and **Prof. Dr. Hilmar Bading** and **Dr. Annarita Patrizi**, thank you for accepting to be part of the committee.

I want to give a special mention to all the current, past and associate members of the Schirmer Lab. Each of you have played an important role during my PhD time. To all the technicians: **Andrea, Ana, katrin** and **Vanesa. Ina**, I am very grateful for your support and for you being by my side through it all, as I do not think I could have finished this crazy ride without you. To all the MD students: **Michael, Sven, Annika, Hannah, Blerina, Thomas** and **Amelie,** thank you for all your help at different moments and for all the happy, stressful and emotional moments that we have shared together. **Patricia**, I have been very lucky to be paired up with you at the GRK. Thank

you for all your help, your spontaneous German lessons and for becoming an essential person in this journey.

To all the PhD members of the lab: **Amel, Julia, Celine** and **Natalie,** and to *our* postdoc **Emilia.** Each of you have been an important support at different times and made the mensa food taste a bit better with your stories and company. Thank you for being kind, always willing to help and becoming my friends along the journey.

To everyone in the lab, I hope the chorus from the ChatGPT song will always stay in our hearts: *'We appeal, we appeal, this decision isn't real.'*

To **Rosanna**, **Rüstem** and **Lukas** for helping me when I was doing my infinite runs of libraries and always occupying your PCRs. I would also like to extend my gratitude to **Simon Hametner** and **Christian Riedl**, for your willingness to always help me and for your warm welcome every time I have visited your lab in Vienna.

To the **Biosciences Faculty of Heidelberg University** for accepting me as a PhD student and **Heidelberg Bioscience International Graduate School (HBIGS)** for supporting me through my doctorate and providing me with new learning opportunities. To the **GRK2727** for funding my PhD and allowing me the opportunity to go to international congresses and lab visits.

To my Heidelberg family **Enrique**, **Sonia** and **Nicola**. You have been there for the darkest of times (DKFZ, first house moving and infinite dramas) but also shared all my successes as your own. They say that hard times bring people closer together, and I could not agree more, since our friendship, even though we move apart, will never fade. **Enrique**, thank you always for being open for mental health walks, playing catan without cheating and for all the talks that we have had that have helped me overcome hard moments. **Sonia**, maybe we are not biological sisters but who cares since you have become one to me. I will always treasure our friendship and your endless energy that made me not be one with the sofa. **Nicola**, my favorite diva. I could not tell you how much I love our gossip sessions and how grateful I am to have gotten to know you; you are truly one of the kindest people I know. Thank you to all three of you, whom without all these years would not have been the same.

To the second heidelberg family members **Cris**, **Ana** and **Sara**. Corona times would have been entirely different without you; our time together felt like a way to escape reality during those challenging days. Thank you for all the parties, hikes, trips and to our annual get togethers, that I

always look forward to. To my lovely two girls from Finland, **Enara** and **Luisa**, who have been following every step of this journey with long calls and get togethers throughout Europe.

I would also like to thank the Saez-Rodriguez lab and Schapiro lab members, since they have been like a second lab family for me. **Julio, Denis, Rico, Daniel, Katrin, Robin, Olga, Sophia, Philipp, Charlotte, Bárbara, Lorna, Jovan, Marja, Chiara, Jan, Martin, Ana, Nico, Taro, Victor, Miguel, Florian** and **Krešo.** I am sure I am missing some of you, but thank you for fostering me when times were rough and welcoming me always as more than a collaborator. To the people I got lucky to meet during my short time at DKFZ but who have stayed there all this time: **Cansu, Nora, Bea, Christina, Daniela, Giulia, Jonas, Paul, Sandro, Vivi** and **Yassin**. I will always remember our gatherings when we were not still stressed to our cores, and when we would meet to keep each other energized through partying. During these nearly 5 years of my stay in Heidelberg, I have had the chance to meet many amazing people that have shaped me into the person I have become, so believe me when I say that I am grateful for having met all of you.

Now I will change to Catalan, the language that I communicate with the people that I love back home.

Primer us voldria donar les gràcies infinites a vosaltres dos, **Papa** i **Mama**. Sense el vostre recolzament durant tot aquest temps sé del cert que no hagués pogut arribar on sóc. Però no només us voldria agraïr aquests últims 5 anys, sinó tota la meva vida. Des de portar-me i recollir-me del cole, donar-me la oportunitat de fer extraescolars, vacances a l'estiu a llocs nous del món, per sempre preocupar-vos per mi i per ser els meus fans número 1. Quan sigui gran, sé que vull ser com vosaltres, tan plena d'energia i vitalitat que faci enrabiar d'enveja a l'altra gent i que sempre tingui la curiositat d' aprendre coses noves. Crec que a l'haver-os tingut de pares això és el que més m'ha marcat, i per això he pogut endinsar-me en una carrera que no és gens fàcil, però és fascinant. Us estimo moltíssim i sempre us estaré agraïda.

Als tetes, **Àlex** i **Guillem**. Sempre, des de petita, heu sigut un exemple a seguir i us admirava moltíssim. De fet, ho continuo fent ja que esteu fent exactament el que voleu i disfrutant les vostres vides al màxim. Quan era més petita sempre em feia una pena gegant el haver-vos de dir adéu quan marxaveu a Alemanya, ja que sentía un buit a dins i no acabava d'entendre perquè ho feieu. Ara que estic en la vostra pell i sé el que se sent, entenc molt millor el vostre punt de vista i tot el que vau haver de sacrificar. Us estimo moltíssim i sempre estaré orgullosa de parlar de vosaltres als demés i de poder dir que sou els meus germans. També voldria agraïr a la

Vanessa i la **Carol**, que em coneixen des de petita i sempre m'han acollit com a una més, i que ara em fan sentir part de les seves noves famílies, a les que tant estimo i em moro de ganes de veure cada cop que torno a casa. **Gràcies als quatre** per haver-me fet tieta dels nens i nenes més guapos del món sencer, i dels que sempre em sento super orgullosa: **Pol, Biel, Júlia, Olivia, Elisa** i **Mar**. Crec que la gent d'Alemanya es pensa que són fins i tot fills meus de les vegades que els parlo d'ells.

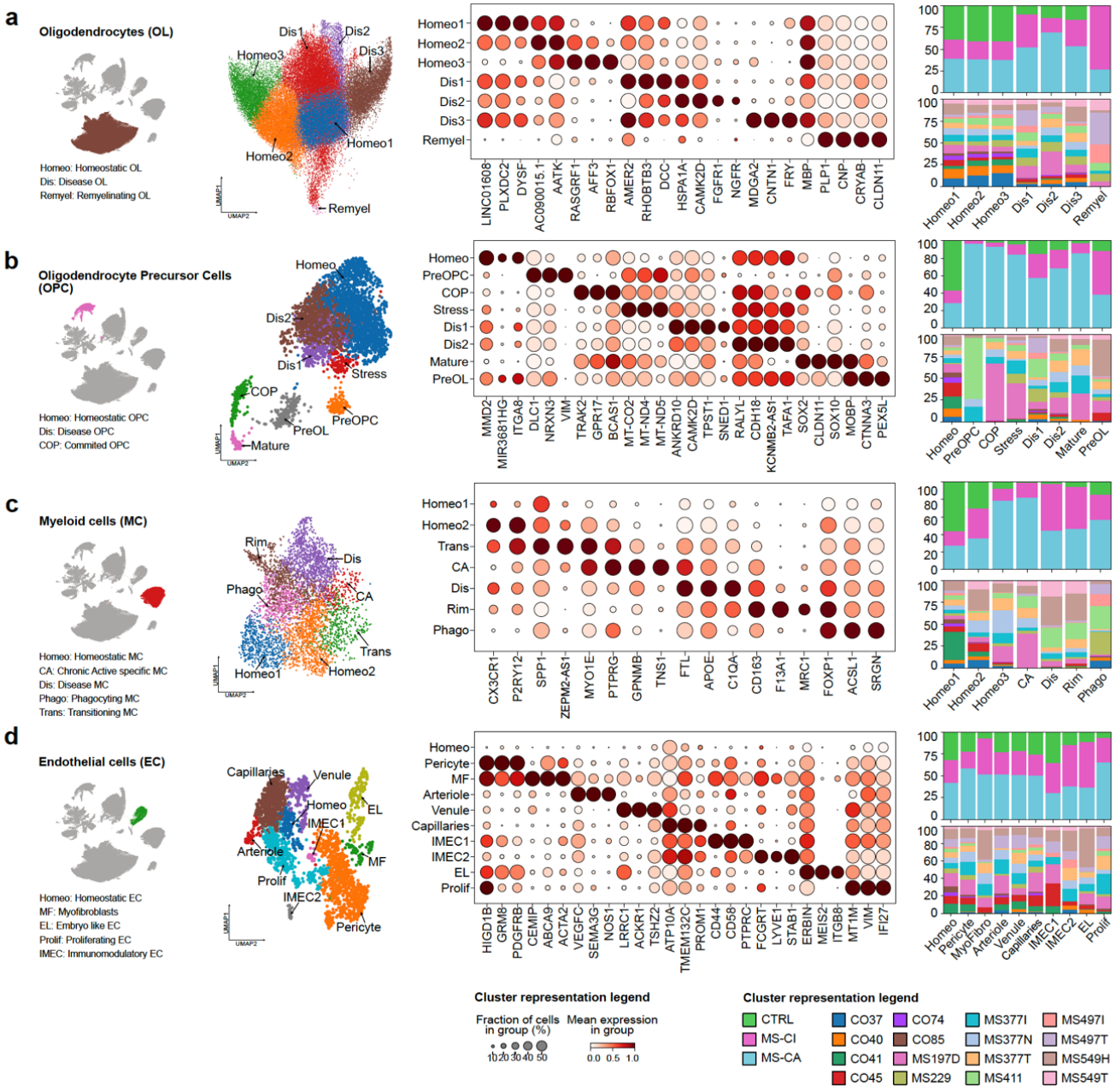
A tú **Pau**, perquè aquest tesis sembla més teva que meva de la de vegades que he escrit el teu nom. Qui ho diria, que quan a la uni no voliem fer treballs junts per si ens enfadavem i ara hem estat treballant colze amb colze en un projecte que ha estat, per posar-ho en poques paraules, intens. Si no fos per tu mai hauria vingut a Alemanya, ni tampoc hagués acabat treballant en aquest grup, que tot i la de drames que hi ha hagut, he après moltíssim. Tu em vas donar la confiança que necessitava per fer aquest pas tan important, i has complert la teva promesa d'acompanyar-me durant tot el trajecte, ja sigui des de fent models estadístics súper complicats (per mi) a tenir trucades per dir-me: 'la coma més a la dreta'. Gràcies per tot, per tenir una paciència infinita, per aguantar-me cada dia (que no és fàcil) i per haver decidit ser el meu company de vida. T'estimo moltíssim.

Als meus amics del cole **Marc** i **Irene**, als que us connec de fa tant de temps, que gairebé sou part de la familia. Gràcies per mantenir el contacte amb mi tot i haver marxat fora i que cada cop que ens veiem sembli que el temps no ha passat.

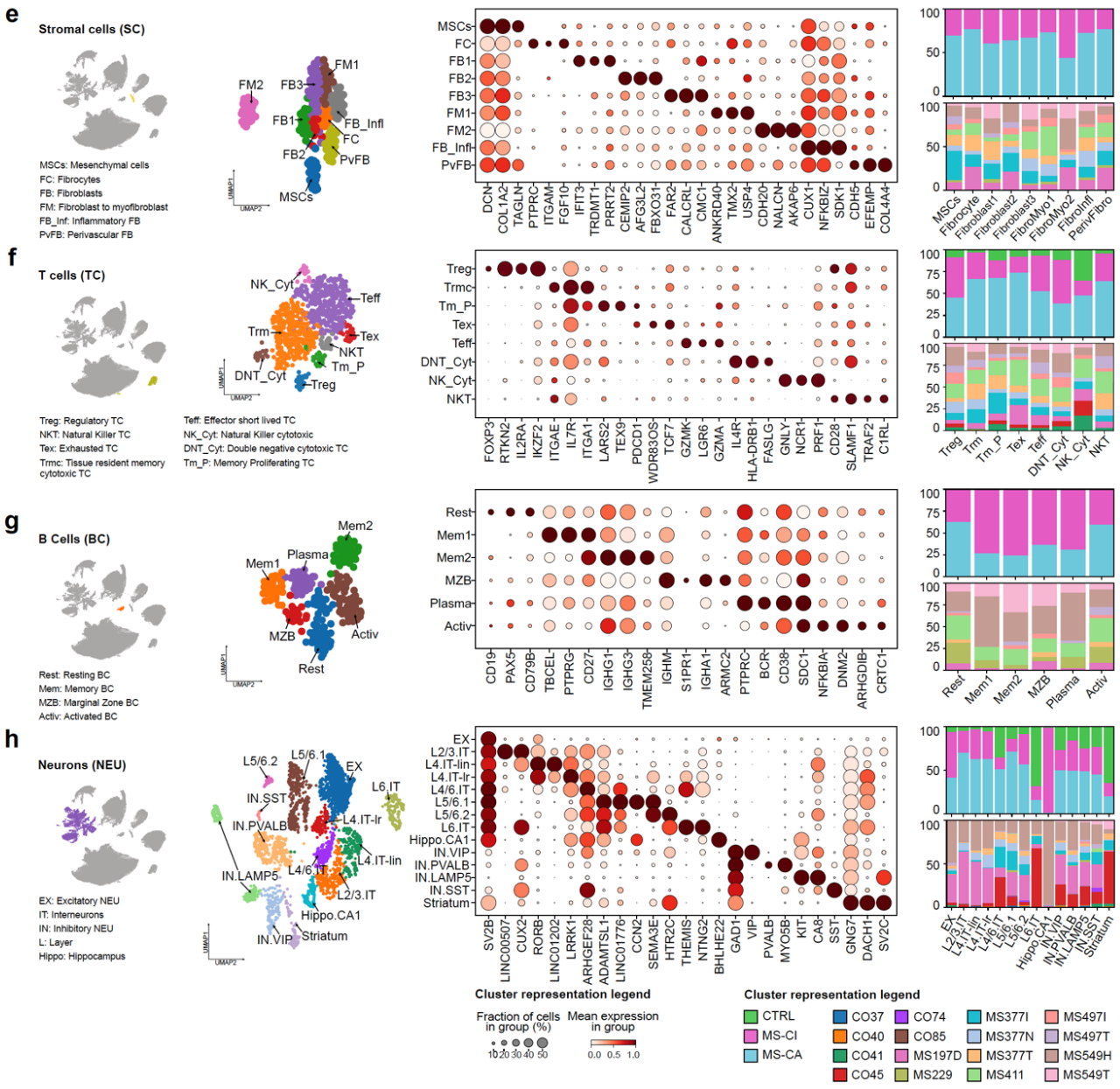
Per últim a mi, per tots els mals moments superats i per tots els reptes que he emprès. Ha sigut una experiència que m'ha convertit en un altre persona, però de la que s'ha d'estar orgullosa.

8. Appendix

Extended Data Figures



Extended Data Figure 1 continues in next page



Extended Data Figure 1. Cell subtype annotation across all main cell types. Repurposed from Lerma-Martin *et al*. 2024. **a,** Oligodendrocytes (OL), **b,** Oligodendrocyte precursor cells (OPC), **c,** Myeloid cells (MC), **d,** Endothelial cells (EC), **e,** Stromal cells (SC), **f,** T cells (TC), **g,** B cells (BC), **h,** Neurons (NEU). For each cell type: original UMAP colored by the given cell type and cell subtype name abbreviations (left); UMAP colored by cell subtypes (center left); dot plot of averaged z-transformed gene expression of marker genes for each cell subtype (center right); proportion of nuclei per condition and snRNA-seq sample grouped by cell subtype (right).

Supplementary tables

Supplementary table 1. Overview of all donor blocks characterized for the study.

Brain Bank ID	Tissue Block	Condition	Lesion Type	Lesions in block	Sex	Age	PMI (hrs)	Duration (years)
MS94	A1D9	MS	CA	1	F	42	11	6
MS197	P2D3	MS	CA	2	F	52	10	NA
MS229	P2C2	MS	CA	1	M	53	13	16
MS371	A3D3	MS	A	1	M	40	27	16
MS371	A3D6	MS	A	1	M	40	27	16
MS377	A2D2	MS	CA	3	F	50	22	23
MS377	A2D4	MS	CA	3	F	50	22	23
MS411	A2A2	MS	CA	2	M	61	24	29
MS426	P2B6	MS	CA	1	F	48	21	29
MS466	A1D6	MS	CI	1	F	65	25	36
MS497	A3C2	MS	CI	2	F	60	26	29
MS513	BSB4	MS	A/CA/CI	5	M	51	17	18
MS528	BSA5	MS	A/CA	2	F	45	17	25
MS530	A1C4	MS	CA	1	M	42	15	21
MS530	P2E5	MS	CA	1	M	42	15	21
MS535	A3A3	MS	CI	1	F	65	12	40
MS535	P2B2	MS	CA	1	F	65	12	40
MS549	A1B2	MS	CA	1	M	50	8	29
MS549	P2D5	MS	CI	2	M	50	8	29
MS584	P2C4	MS	A	1	F	42	36	12
MS586	A1B2	MS	A	1	F	57	27	20
PDCO23	P2C4	Control	NA	0	F	78	23	NA
PDCO40	A1B2	Control	NA	0	F	61	15	NA
CO37	P5B3	Control	NA	0	M	87	5	NA
CO41	A1C4	Control	NA	0	M	56	20	NA
CO45	A1D4	Control	NA	0	M	77	22	NA
CO85	A3C2	Control	NA	0	F	81	22	NA
CO74	A1A2	Control	NA	0	F	84	22	NA
CO96	A3C3	Control	NA	0	F	68	NA	NA

Supplementary table 2. Overview samples used for transcriptomics.

Original ID	Given ID	Lesion	Sex	RIN	snRNA-seq	Batch	ST	Batch
MS94 A1D9	MS94	CA	F	8.7	No	-	Yes	1
MS197 P2D3	MS197U	CA	F	9	Yes	4	Yes	1
MS197 P2D3	MS197D	CA	F	9	No	4	Yes	1
MS229 P2C2	MS229	CA	M	7	Yes	4	Yes	1
MS377 A2D2	MS377N	CA	F	8.9	Yes	4	Yes	1
MS377 A2D4	MS377I	CA	F	6.5	Yes	1	Yes	1
MS377 A2D4	MS377T	CA	F	6.5	Yes	1	Yes	1
MS411 A2A2	MS411	CA	M	5.9	Yes	1	Yes	1
MS497 A3C2	MS497I	CI	F	6.1	Yes	3	Yes	1
MS497 A3C2	MS497T	CI	F	6.1	Yes	2	Yes	1
MS549 P2D5	MS549H	CI	M	8	Yes	2	Yes	1
MS549 P2D5	MS549T	CI	M	8	Yes	3	Yes	1
CO37 P5B3	CO37	Ctrl	M	6.1	Yes	4	Yes	3
PDCO40 A1B2	CO40	Ctrl	F	7.8	Yes	1	Yes	1
CO41 A1C4	CO41	Ctrl	M	7.2	Yes	4	Yes	3
CO45 A1D4	CO45	Ctrl	M	6.4	Yes	4	No	-
CO74 A1A2	CO74	Ctrl	F	7	Yes	2	Yes	2
CO85 A3C2	CO85	Ctrl	F	5.9	Yes	3	Yes	1
CO96 A3C3	CO96	Ctrl	F	7.7	No	-	Yes	1

Supplementary table 3. Predicted cell-cell communication events in CTRL

Interaction	Source_cs	Target_cs
OPC (NLGN3) - OL (NRXN3)	-	OL_Homeo1, OL_Homeo3
OL (NLGN3) - OL (NRXN3)	-	OL_Homeo1, OL_Homeo3
OL (NRXN3) - OPC (NLGN3)	OL_Homeo1, OL_Homeo3	-
OL (NRXN3) - OL (NLGN3)	OL_Homeo1, OL_Homeo3	-
OPC (NLGN3) - OL (NRXN2)	-	OL_Homeo2, OL_Homeo3
OL (NLGN3) - OL (NRXN2)	-	OL_Homeo2, OL_Homeo3

OL (NRXN2) - OPC (NLGN3)	OL_Homeo2, OL_Homeo3	-
OL (NRXN2) - OL (NLGN3)	OL_Homeo2, OL_Homeo3	-
OPC (TGFA) - OL (ERBB3)	OPC_Homeo	-
OL (TGFA) - OL (ERBB3)	OL_Homeo3	-
OPC (LTBP1) - OL (ITGB5)	-	OL_Homeo2, OL_Homeo3
OL (LTBP1) - OL (ITGB5)	OL_Homeo1	OL_Homeo2, OL_Homeo3
OPC (NCAM1) - OL (FGFR2)	-	OL_Homeo3
OPC (NCAM1) - OPC (FGFR1)	-	OPC_Homeo
OPC (NCAM1) - AS (GFRA1)	-	AS_Homeo
OPC (VEGFA) - OL (TYRO3)	OPC_Homeo	-
OPC (NTN1) - OL (UNC5B)	-	OL_Homeo2, OL_Homeo3
OPC (LPL) - EC (GPIHBP1)	OPC_Homeo	-
OPC (VEGFA) - MC (FLT1)	OPC_Homeo	-
OPC (VEGFA) - MC (ITGAV)	OPC_Homeo	-
OPC (VEGFA) - MC (EPHB2)	OPC_Homeo	MC_Homeo2
AS (VEGFA) - MC (EPHB2)	AS_Homeo	MC_Homeo2
EC (VEGFA) - MC (EPHB2)	-	MC_Homeo2
OPC (VEGFA) - OL (EPHB2)	OPC_Homeo	OL_Homeo2, OL_Homeo3
AS (VEGFA) - OL (EPHB2)	AS_Homeo	OL_Homeo2, OL_Homeo3
EC (VEGFA) - OL (EPHB2)	-	OL_Homeo2, OL_Homeo3
OPC (VEGFA) - OL (SIRPA)	OPC_Homeo	OL_Homeo3
EC (VEGFA) - OL (SIRPA)	-	OL_Homeo3

OPC (VEGFA) - MC (SIRPA)	OPC_Homeo	-
OPC (RIMS2) - OL (ABCA1)	OPC_Homeo	OL_Homeo1, OL_Homeo2, OL_Homeo3
OPC (NTN1) - OL (MCAM)	-	OL_Homeo1
OL (APP) - OL (VLDLR)	OL_Homeo1, OL_Homeo2, OL_Homeo3	OL_Homeo2, OL_Homeo3
OL (APP) - EC (TSPAN15)	OL_Homeo1, OL_Homeo2, OL_Homeo3	-
OL (APP) - OL (TSPAN15)	OL_Homeo1, OL_Homeo2, OL_Homeo3	-
OL (APP) - OL (LRP6)	OL_Homeo1, OL_Homeo2, OL_Homeo3	OL_Homeo1
OL (SEMA3B) - OPC (NRP2)	OL_Homeo2, OL_Homeo3	-
OL (SEMA6A) - OPC (PLXNA2)	OL_Homeo3	-
AS (SEMA6A) - OPC (PLXNA2)	AS_Homeo	-
OL (FARP2) - OPC (PLXNA2)	OL_Homeo3	-
OL (FARP2) - OPC (PLXNA4)	OL_Homeo3	-
OL (SEMA3B) - OPC (NRP1)	OL_Homeo2, OL_Homeo3	-
OL (SEMA4D) - MC (PLXNB2)	OL_Homeo2, OL_Homeo3	-
OL (SEMA6A) - OPC (PLXNA4)	OL_Homeo3	-
AS (SEMA6A) - OPC (PLXNA4)	AS_Homeo	-
OL (FGF1) - OL (FGFR2)	OL_Homeo3	OL_Homeo3
OL (FGF1) - OPC (FGFR3)	OL_Homeo3	OPC_Homeo
OL (FGF1) - OPC (FGFR1)	OL_Homeo3	OPC_Homeo
OL (FGF1) - EC (FGFR1)	OL_Homeo3	-
OL (PSEN1) - OL (NOTCH1)	OL_Homeo2, OL_Homeo3	-
OL (PSEN1) - MC (NOTCH1)	OL_Homeo2, OL_Homeo3	-

OL (PSEN1) - AS (NOTCH1)	OL_Homeo2, OL_Homeo3	AS_Homeo
OL (PSEN1) - EC (NCSTN)	OL_Homeo2, OL_Homeo3	-
OL (BMP8A) - OPC (ACVR1)	OL_Homeo2, OL_Homeo3	-
OL (BMP8A) - AS (ACVR1)	OL_Homeo2, OL_Homeo3	-
OL (BMP8A) - MC (TGFBR1)	OL_Homeo2, OL_Homeo3	-
OL (CGN) - MC (TGFBR1)	OL_Homeo3	-
OL (HLA-DMA) - MC (CD4)	OL_Homeo2, OL_Homeo3	-
OL (TGFB3) - OL (ITGB5)	-	OL_Homeo2, OL_Homeo3
OL (LRRC4B) - OL (PTPRS)	OL_Homeo2, OL_Homeo3	-
OL (CDH1) - OL (ERBB3)	OL_Homeo1, OL_Homeo2, OL_Homeo3	-
OL (CDH1) - OPC (PTPRM)	OL_Homeo1, OL_Homeo2, OL_Homeo3	-
OL (CDH1) - AS (PTPRM)	OL_Homeo1, OL_Homeo2, OL_Homeo3	-
MC (ST6GAL1) - OL (CD22)	MC_Homeo2	OL_Homeo3
MC (TGFB1) - OL (ITGB5)	-	OL_Homeo2, OL_Homeo3
MC (COL18A1) - OL (ITGB5)	MC_Homeo2	OL_Homeo2, OL_Homeo3
MC (COL18A1) - OL (PTPRS)	MC_Homeo2	-
MC (ADAM17) - MC (RHBDF2)	-	MC_Homeo2
MC (SPP1) - MC (ITGAV)	MC_Homeo2	-
AS (BMP7) - OPC (ACVR1)	AS_Homeo	-
AS (ANGPTL4) - OL (CDH11)	AS_Homeo	OL_Homeo2, OL_Homeo3
AS (NTRK3) - OL (PTPRS)	AS_Homeo	-
AS (NID1) - MC (ITGAV)	AS_Homeo	-

AS (SPON1) - MC (LRP8)	AS_Homeo	MC_Homeo2
AS (SPON1) - OPC (LRP8)	AS_Homeo	OPC_Homeo
AS (AFDN) - OL (EPHB2)	-	OL_Homeo2, OL_Homeo3
EC (AFDN) - OL (EPHB2)	-	OL_Homeo2, OL_Homeo3
AS (AFDN) - MC (EPHB2)	-	MC_Homeo2
EC (AFDN) - MC (EPHB2)	-	MC_Homeo2
EC (BSG) - OL (SLC16A1)	-	OL_Homeo1, OL_Homeo3
EC (NECTIN2) - OL (CD226)	-	OL_Homeo3
MC (NECTIN2) - OL (CD226)	-	OL_Homeo3
EC (LGALS9) - MC (MRC2)	-	MC_Homeo2

Supplementary table 4. Predicted cell-cell communication events in MS-CA lesions

Interaction	Source_cs	Target_cs
AS (HMGB1) - MC (CD163)	AS_Dis1	MC_CA, MC_Dis, MC_Rim
EC (HMGB1) - MC (CD163)	EC_prolif	MC_CA, MC_Dis, MC_Rim
OL (HMGB1) - MC (CD163)	-	MC_CA, MC_Dis, MC_Rim
AS (HMGB1) - TC (CXCR4)	AS_Dis1	-
AS (HMGB1) - BC (CXCR4)	AS_Dis1	-
AS (HMGB1) - MC (TLR2)	AS_Dis1	MC_Trans, MC_Rim
TC (HMGB1) - MC (TLR2)	-	MC_Trans, MC_Rim
EC (HMGB1) - MC (TLR2)	EC_Prolif	MC_Trans, MC_Rim
AS (HMGB1) - BC (SDC1)	AS_Dis1	-
AS (LAMC1) - AS (ITGA3)	AS_Dis1	AS_Dis1
AS (VEGFD) - MC (FLT1)	AS_Dis1	MC_CA, MC_Trans

AS (VEGFD) MC (ITGA4)	AS_Dis1	MC_Dis, MC_Rim
OPC (ADAM23) TC (ITGA4)	-	TC_Effector
AS (VCAN) - TC (ITGA4)	AS_R	TC_Effector
OPC (VCAN) - TC (ITGA4)	-	TC_Effector
AS (VCAN) - TC (CD44)	AS_R	-
AS (LAMA2) - TC (CD44)	AS_Dis1	-
AS (FARP2) MC (PLXNA1)	-	MC_CA
OL (FARP2) MC (PLXNA1)	OL_Dis2	MC_CA
AS (SEMA6D) - AS (PLXNA1)	-	AS_Dis1
AS (SEMA5B) - MC (PLXNA1)	-	MC_CA
AS (SEMA5B) - AS (PLXNA1)	-	AS_Dis1
OPC (SEMA5B) - MC (PLXNA1)	-	MC_CA
OPC (SEMA5B) - AS (PLXNA1)	-	AS_Dis1
AS (PAM) - AS (FAP)	AS_Dis1, AS_Dis2	AS_Dis1, AS_Dis2, AS_R
AS (BGN) - MC (TLR2)	AS_Dis1, AS_R	MC_Trans, MC_Rim
AS (ANXA1) - MC (FPR1)	AS_Dis1, AS_Dis2, AS_R	MC_Dis, MC_Trans
AS (CSF1) - MC (SIRPA)	AS_R	MC_CA, MC_Trans
OL (CSF1) - MC (SIRPA)	OL_Dis1	MC_CA, MC_Trans
AS (COL8A1) - MC (ITGA1)	AS_Dis1, AS_Dis2	MC_Trans
AS (COL8A1) - AS (ITGA10)	AS_Dis1, AS_Dis2	-
AS (COL8A1) - EC (ITGA10)	AS_Dis1, AS_Dis2	-
AS (COL8A1) - MC (ITGA11)	AS_Dis1, AS_Dis2	-
AS (COL24A1) - MC (ITGA1)	AS_Dis1	MC_Trans
AS (COL24A1) AS (ITGA10)	AS_Dis1	-
AS (COL24A1) OPC (ITGA10)	AS_Dis1	-

AS (COL24A1) EC (ITGA10)	AS_Dis1	-
AS (COL24A1) - MC (ITGA11)	AS_Dis1	-
AS (COL24A1) - OPC (ITGA11)	AS_Dis1	-
AS (COL27A1) - MC (ITGA1)	AS_Dis1, AS_Dis2	MC_Trans
AS (COL27A1) - MC (ITGA11)	AS_Dis1, AS_Dis2	-
AS (ANGPT1) - MC (ITGB1)	AS_Dis1	MC_Dis
AS (CADM3) - MC (CADM1)	AS_Dis2, AS_R	MC_CA, MC_Dis, MC_Rim
AS (C3) - MC (CD81)	AS_Dis1, AS_Dis2, AS_R	MC_Dis
AS (EGF) - AS (TNFRSF11B)	AS_Dis1, AS_Dis2, AS_R	AS_Dis1, AS_R
AS (SORBS1) - BC (INSR)	AS_Dis1	-
AS (HSP90B1) - MC (TLR1)	AS_Dis1, AS_R	MC_Trans
MC (SERPING1) - AS (LRP1)	MC_Dis, MC_Trans	AS_Dis1, AS_Dis2, AS_R
MC (C1QB) - AS (LRP1)	MC_Dis	AS_Dis1, AS_Dis2, AS_R
MC (APOC2) - AS (LRP1)	MC_Dis, MC_Trans	AS_Dis1, AS_Dis2, AS_R
MC (ANXA2) - MC (TLR2)	MC_Trans, MC_Dis, MS_CA	MC_Trans, MC_Rim
AS (ANXA2) - MC (TLR2)	AS_Dis1, AS_Dis2, AS_R	MC_Trans, MC_Rim
MC (LGALS9) - TC (PTPRC)	MC_Trans, MC_Dis, MC_CA	-
MC (TNFSF13) - AS (TNFRSF1A)	MC_Dis, MC_Rim	AS_Dis1, AS_Dis2, AS_R
MC (TNFSF13) - AS (TNFRSF11B)	MC_Dis, MC_Rim	AS_Dis1, AS_R
MC (C5) - AS (C5AR1)	MC_Trans, MC_Dis, MC_CA, MC_Rim	
MC (RPS19) - AS (C5AR1)	MC_Dis	-
EC (RPS19) - AS (C5AR1)	EC_Prolif	
MC (FGL1) - OPC (EGFR)	MC_Trans	-
MC (CALR) - AS (ITGA3)	MC_Dis	AS_Dis1
MC (DSCAM) - BC (DCC)	MC_Trans	-

MC (APOE) - TC (SORL1)	MC_Dis, MC_Rim	-
MC (B2M) - TC (CD247)	MC_Dis, MC_Rim	-
EC (B2M) - TC (CD247)	EC_Prolif	-
MC (SPP1) - TC (CD44)	MC_Trans	-
MC (SPP1) - BC (CD44)	MC_Trans	-
MC (SPP1) - TC (ITGA4)	MC_Trans	TC_Effector
MC (ADAM28) - TC (ITGA4)	-	TC_Effector
MC (CD14) - MC (ITGA4)	MC_Dis, MC_Rim	MC_Dis, MC_Rim
MC (CD14) - MC (ITGB1)	MC_Dis, MC_Rim	MC_Dis
MC (CD14) - AS (ITGB1)	MC_Dis, MC_Rim	AS_Dis1, AS_R
MC (CD14) - EC (ITGB1)	MC_Dis, MC_Rim	EC_Prolif
MC (CD14) - OL (RIPK1)	MC_Dis, MC_Rim	OL_Dis2
MC (CD14) - MC (TLR1)	MC_Dis, MC_Rim	MC_Trans
EC (VIM) - TC (CD44)	EC_prolif	-
TC (VIM) - AS (CD44)	-	AS_Dis1, AS_Dis2
EC (HLA-E) - TC (KLRK1)	EC_prolif	TC_Effector
EC (HLA-E) - TC (CD8A)	EC_prolif	TC_Effector
EC (CD59) - TC (CD2)	EC_prolif	TC_Effector, TC_Trm
EC (TGFB3) - MC (ENG)	-	MC_Trans
EC (TGM2) - BC (ITGA4)	EC_prolif	-
EC (TGM2) - TC (ITGA4)	EC_prolif	TC_Effector
EC (VWF) - BC (ITGA4)	EC_prolif	-
EC (VWF) - TC (ITGA4)	EC_prolif	TC_Effector
EC (SORBS1) - TC (ITGA1)	-	TC_Trm
SC (SORBS1) - TC (ITGA1)	-	TC_Trm

SC (LAMA2) - TC (ITGA1)	-	TC_Trm
EC (B2M) - BC (LILRB1)	EC_prolif	
EC (B2M) - TC (KLRD1)	EC_prolif	-
EC (TFPI) - AS (LRP1)	-	AS_Dis1, AS_Dis2, AS_R
TC (MAML2) - MC (NOTCH2)	-	MC_CA, MC_Dis, MC_Rim
TC (CCL5) - OPC (GRM7)	TC_Effector	-
BC (CD38) - EC (PECAM1)	-	EC_Prolif
BC (GAS6) - EC (MERTK)	-	MC_Trans, MC_CA
AS (GAS6) - MC (MERTK)	AS_Dis2, AS_R	MC_Trans, MC_CA
BC (HSP90AA1) - AS (EGFR)	-	AS_Dis2
EC (HSP90AA1) - OPC (EGFR)	EC_Prolif	-
BC (HSP90B1) - MC (TLR2)	-	MC_Trans, MC_Rim
EC (HSP90B1) - MC (TLR2)	EC_Prolif	MC_Trans, MC_Rim
AS (HSP90B1) - MC (TLR2)	AS_Dis, AS_R	MC_Trans, MC_Rim
BC (HSP90B1) - MC (TLR1)	-	MC_Trans
BC (HSP90B1) - AS (LRP1)	-	AS_Dis1, AS_Dis2, AS_R
BC (HSP90B1) - MC (LRP1)	-	MC_CA, MC_Dis, MC_Trans
MC (HSP90B1) - AS (LRP1)	MC_Dis, MC_Trans	AS_Dis1, AS_Dis2, AS_R
BC (CALM1) - AS (KCNQ5)	-	AS_Dis2
BC (CALM1) - AS (CACNA1C)	-	AS_Dis1
BC (COL4A4) - AS (CD44)	-	AS_Dis1, AS_Dis2
BC (ADAM10) - AS (CD44)	-	AS_Dis1, AS_Dis2
BC (ADAM10) - AS (NRCAM)	-	AS_Dis1, AS_Dis2
BC (ST6GAL1) - AS (EGFR)	-	AS_Dis2

Supplementary table table 5. Predicted cell-cell communication events in MS-CI lesions

Interaction	Source_cs	Target_cs
AS (EGF) - AS (TNFRSF11B)	AS_R	AS_R
AS (VCAN) - BC (CD44)	AS_R	-

9. References

1. Lassmann, H. Multiple Sclerosis Pathology. *Cold Spring Harb. Perspect. Med.* **8**, (2018).
2. Gold, R., Linington, C. & Lassmann, H. Understanding pathogenesis and therapy of multiple sclerosis via animal models: 70 years of merits and culprits in experimental autoimmune encephalomyelitis research. *Brain* **129**, 1953–1971 (2006).
3. Zamecnik, C. R. *et al.* An autoantibody signature predictive for multiple sclerosis. *Nat. Med.* **30**, 1300–1308 (2024).
4. Häusser-Kinzel, S. & Weber, M. S. The role of B cells and antibodies in multiple sclerosis, neuromyelitis optica, and related disorders. *Front. Immunol.* **10**, 201 (2019).
5. Haase, S. & Linker, R. A. Inflammation in multiple sclerosis. *Ther. Adv. Neurol. Disord.* **14**, 17562864211007687 (2021).
6. Lassmann, H. Pathogenic Mechanisms Associated With Different Clinical Courses of Multiple Sclerosis. *Front. Immunol.* **9**, 3116 (2018).
7. Monaco, S., Nicholas, R., Reynolds, R. & Magliozzi, R. Intrathecal Inflammation in Progressive Multiple Sclerosis. *Int. J. Mol. Sci.* **21**, (2020).
8. Reich, D. S., Lucchinetti, C. F. & Calabresi, P. A. Multiple Sclerosis. *N. Engl. J. Med.* **378**, 169–180 (2018).
9. Kuhlmann, T. *et al.* An updated histological classification system for multiple sclerosis lesions. *Acta Neuropathol.* **133**, 13–24 (2017).
10. Lucchinetti, C. *et al.* Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. *Ann. Neurol.* **47**, 707–717 (2000).
11. Ascherio, A., Munger, K. L. & Simon, K. C. Vitamin D and multiple sclerosis. *Lancet Neurol.* **9**, 599–612 (2010).
12. Colotta, F., Jansson, B. & Bonelli, F. Modulation of inflammatory and immune responses by vitamin D. *J. Autoimmun.* **85**, 78–97 (2017).

13. Charoenngam, N. & Holick, M. F. Immunologic Effects of Vitamin D on Human Health and Disease. *Nutrients* **12**, (2020).
14. Munger, K. L. *et al.* Childhood body mass index and multiple sclerosis risk: a long-term cohort study. *Mult. Scler.* **19**, 1323–1329 (2013).
15. Ascherio, A. & Munger, K. L. Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. *Ann. Neurol.* **61**, 504–513 (2007).
16. Ascherio, A. & Munger, K. L. Epstein-barr virus infection and multiple sclerosis: a review. *J. Neuroimmune Pharmacol.* **5**, 271–277 (2010).
17. Pender, M. P. Infection of autoreactive B lymphocytes with EBV, causing chronic autoimmune diseases. *Trends Immunol.* **24**, 584–588 (2003).
18. Kale, N. Optic neuritis as an early sign of multiple sclerosis. *Eye Brain* **8**, 195–202 (2016).
19. Ciapă, M. A., Șalaru, D. L., Stătescu, C., Sascău, R. A. & Bogdănici, C. M. Optic Neuritis in Multiple Sclerosis-A Review of Molecular Mechanisms Involved in the Degenerative Process. *Curr. Issues Mol. Biol.* **44**, 3959–3979 (2022).
20. Wu, G. F., Parker Harp, C. R. & Shindler, K. S. Optic Neuritis: A Model for the Immuno-pathogenesis of Central Nervous System Inflammatory Demyelinating Diseases. *Curr. Immunol. Rev.* **11**, 85–92 (2015).
21. Abou Zeid, N. & Bhatti, M. T. Acute inflammatory demyelinating optic neuritis: evidence-based visual and neurological considerations. *Neurologist* **14**, 207–223 (2008).
22. Beck, R. W. & Gal, R. L. Treatment of acute optic neuritis: a summary of findings from the optic neuritis treatment trial. *Arch. Ophthalmol.* **126**, 994–995 (2008).
23. Optic Neuritis Study Group. Multiple sclerosis risk after optic neuritis: final optic neuritis treatment trial follow-up. *Arch. Neurol.* **65**, 727–732 (2008).
24. Balcer, L. J. Clinical practice. Optic neuritis. *N. Engl. J. Med.* **354**, 1273–1280 (2006).
25. Harrow-Mortelliti, M., Reddy, V. & Jimsheleishvili, G. Physiology, Spinal Cord. in *StatPearls [Internet]* (StatPearls Publishing, 2023).

26. Ouellette, R. *et al.* 7 T imaging reveals a gradient in spinal cord lesion distribution in multiple sclerosis. *Brain* **143**, 2973–2987 (2020).
27. Bartanusz, V., Jezova, D., Alajajian, B. & Digicaylioglu, M. The blood-spinal cord barrier: morphology and clinical implications. *Ann. Neurol.* **70**, 194–206 (2011).
28. Gilmore, C. P. *et al.* Spinal cord grey matter lesions in multiple sclerosis detected by post-mortem high field MR imaging. *Mult. Scler.* **15**, 180–188 (2009).
29. Bot, J. C. J. *et al.* Spinal cord abnormalities in recently diagnosed MS patients: added value of spinal MRI examination. *Neurology* **62**, 226–233 (2004).
30. Göçmen, R. The Relevance of Neuroimaging Findings to Physical Disability in Multiple Sclerosis. *Noro Psikiyatr Ars* **55**, S31–S36 (2018).
31. DeLuca, G. C., Williams, K., Evangelou, N., Ebers, G. C. & Esiri, M. M. The contribution of demyelination to axonal loss in multiple sclerosis. *Brain* **129**, 1507–1516 (2006).
32. Kearney, H., Miller, D. H. & Ciccarelli, O. Spinal cord MRI in multiple sclerosis--diagnostic, prognostic and clinical value. *Nat. Rev. Neurol.* **11**, 327–338 (2015).
33. Cohen, A. B. *et al.* The relationships among MRI-defined spinal cord involvement, brain involvement, and disability in multiple sclerosis. *J. Neuroimaging* **22**, 122–128 (2012).
34. Kreiter, D., Postma, A. A., Hupperts, R. & Gerlach, O. Hallmarks of spinal cord pathology in multiple sclerosis. *J. Neurol. Sci.* **456**, 122846 (2024).
35. Nakashima, I., Fujihara, K., Okita, N., Takase, S. & Itoyama, Y. Clinical and MRI study of brain stem and cerebellar involvement in Japanese patients with multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **67**, 153–157 (1999).
36. Yang, Y. *et al.* Cerebellar and/or Brainstem Lesions Indicate Poor Prognosis in Multiple Sclerosis: A Systematic Review. *Front. Neurol.* **13**, 874388 (2022).
37. Wilkins, A. Cerebellar Dysfunction in Multiple Sclerosis. *Front. Neurol.* **8**, 312 (2017).
38. Fritz, N. E. *et al.* Cerebellar Contributions to Motor and Cognitive Control in Multiple Sclerosis. *Arch. Phys. Med. Rehabil.* **103**, 1592–1599 (2022).

39. Bø, L., Vedeler, C. A., Nyland, H. I., Trapp, B. D. & Mørk, S. J. Subpial demyelination in the cerebral cortex of multiple sclerosis patients. *J. Neuropathol. Exp. Neurol.* **62**, 723–732 (2003).
40. Kutzelnigg, A. *et al.* Cortical demyelination and diffuse white matter injury in multiple sclerosis. *Brain* **128**, 2705–2712 (2005).
41. Lucchinetti, C. F. *et al.* Inflammatory cortical demyelination in early multiple sclerosis. *N. Engl. J. Med.* **365**, 2188–2197 (2011).
42. Peterson, J. W., Bö, L., Mörk, S., Chang, A. & Trapp, B. D. Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. *Ann. Neurol.* **50**, 389–400 (2001).
43. Bö, L., Geurts, J. J. G., van der Valk, P., Polman, C. & Barkhof, F. Lack of correlation between cortical demyelination and white matter pathologic changes in multiple sclerosis. *Arch. Neurol.* **64**, 76–80 (2007).
44. Popescu, B. F. G. & Lucchinetti, C. F. Meningeal and cortical grey matter pathology in multiple sclerosis. *BMC Neurol.* **12**, 11 (2012).
45. Bö, L., Geurts, J. J. G., Mörk, S. J. & van der Valk, P. Grey matter pathology in multiple sclerosis. *Acta Neurol. Scand. Suppl.* **183**, 48–50 (2006).
46. Moll, N. M. *et al.* Cortical demyelination in PML and MS: Similarities and differences. *Neurology* **70**, 336–343 (2008).
47. Kidd, D. *et al.* Cortical lesions in multiple sclerosis. *Brain* **122 (Pt 1)**, 17–26 (1999).
48. Serafini, B., Rosicarelli, B., Magliozzi, R., Stigliano, E. & Aloisi, F. Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis. *Brain Pathol.* **14**, 164–174 (2004).
49. Franciotta, D., Salvetti, M., Lolli, F., Serafini, B. & Aloisi, F. B cells and multiple sclerosis. *Lancet Neurol.* **7**, 852–858 (2008).

50. Pikor, N. B., Prat, A., Bar-Or, A. & Gommerman, J. L. Meningeal Tertiary Lymphoid Tissues and Multiple Sclerosis: A Gathering Place for Diverse Types of Immune Cells during CNS Autoimmunity. *Front. Immunol.* **6**, 657 (2015).
51. Serafini, B. *et al.* Dysregulated Epstein-Barr virus infection in the multiple sclerosis brain. *J. Exp. Med.* **204**, 2899–2912 (2007).
52. Absinta, M. *et al.* Gadolinium-based MRI characterization of leptomeningeal inflammation in multiple sclerosis. *Neurology* **85**, 18–28 (2015).
53. Magliozzi, R. *et al.* Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology. *Brain* **130**, 1089–1104 (2007).
54. Ludwin, S. K., Rao, V. T., Moore, C. S. & Antel, J. P. Astrocytes in multiple sclerosis. *Mult. Scler.* **22**, 1114–1124 (2016).
55. Steinman, L. Multiple sclerosis: a coordinated immunological attack against myelin in the central nervous system. *Cell* **85**, 299–302 (1996).
56. Bagasra, O. *et al.* Activation of the inducible form of nitric oxide synthase in the brains of patients with multiple sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 12041–12045 (1995).
57. Mycko, M. P., Papoian, R., Boschert, U., Raine, C. S. & Selmaj, K. W. cDNA microarray analysis in multiple sclerosis lesions: detection of genes associated with disease activity. *Brain* **126**, (2003).
58. Holman, D. W., Klein, R. S. & Ransohoff, R. M. The blood-brain barrier, chemokines and multiple sclerosis. *Biochim. Biophys. Acta* **1812**, 220–230 (2011).
59. Larochelle, C. *et al.* Melanoma cell adhesion molecule identifies encephalitogenic T lymphocytes and promotes their recruitment to the central nervous system. *Brain* **135**, 2906–2924 (2012).
60. Larochelle, C. *et al.* Melanoma cell adhesion molecule-positive CD8 T lymphocytes mediate central nervous system inflammation. *Ann. Neurol.* **78**, 39–53 (2015).

61. Ifergan, I. *et al.* Role of Ninjurin-1 in the migration of myeloid cells to central nervous system inflammatory lesions. *Ann. Neurol.* **70**, 751–763 (2011).
62. Mishra, M. K. & Yong, V. W. Myeloid cells - targets of medication in multiple sclerosis. *Nat. Rev. Neurol.* **12**, 539–551 (2016).
63. Aguzzi, A., Barres, B. A. & Bennett, M. L. Microglia: scapegoat, saboteur, or something else? *Science* **339**, 156–161 (2013).
64. Ruprecht, K. *et al.* Effects of oncostatin M on human cerebral endothelial cells and expression in inflammatory brain lesions. *J. Neuropathol. Exp. Neurol.* **60**, 1087–1098 (2001).
65. Henderson, A. P. D., Barnett, M. H., Parratt, J. D. E. & Prineas, J. W. Multiple sclerosis: distribution of inflammatory cells in newly forming lesions. *Ann. Neurol.* **66**, 739–753 (2009).
66. Perry, V. H. A revised view of the central nervous system microenvironment and major histocompatibility complex class II antigen presentation. *J. Neuroimmunol.* **90**, 113–121 (1998).
67. Wu, C. *et al.* Sialoadhesin-positive macrophages bind regulatory T cells, negatively controlling their expansion and autoimmune disease progression. *J. Immunol.* **182**, 6508–6516 (2009).
68. Cantuti-Castelvetri, L. *et al.* Defective cholesterol clearance limits remyelination in the aged central nervous system. *Science* **359**, 684–688 (2018).
69. Miron, V. E. *et al.* M2 microglia and macrophages drive oligodendrocyte differentiation during CNS remyelination. *Nat. Neurosci.* **16**, 1211–1218 (2013).
70. Stys, P. K., Zamponi, G. W., van Minnen, J. & Geurts, J. J. G. Will the real multiple sclerosis please stand up? *Nat. Rev. Neurosci.* **13**, 507–514 (2012).
71. Marik, C., Felts, P. A., Bauer, J., Lassmann, H. & Smith, K. J. Lesion genesis in a subset of patients with multiple sclerosis: a role for innate immunity? *Brain* **130**, 2800–2815 (2007).
72. Lassmann, H. Mechanisms of white matter damage in multiple sclerosis. *Glia* **62**, 1816–1830 (2014).
73. Li, R. *et al.* Proinflammatory GM-CSF-producing B cells in multiple sclerosis and B cell depletion therapy. *Sci. Transl. Med.* **7**, 310ra166 (2015).

74. Prineas, J. W. *et al.* Immunopathology of secondary-progressive multiple sclerosis. *Ann. Neurol.* **50**, 646–657 (2001).
75. Frischer, J. M. *et al.* The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain* **132**, 1175–1189 (2009).
76. Hametner, S., Dal Bianco, A., Trattnig, S. & Lassmann, H. Iron related changes in MS lesions and their validity to characterize MS lesion types and dynamics with Ultra-high field magnetic resonance imaging. *Brain Pathol.* **28**, 743–749 (2018).
77. Tham, M. *et al.* Iron Heterogeneity in Early Active Multiple Sclerosis Lesions. *Ann. Neurol.* **89**, 498–510 (2021).
78. Goldschmidt, T., Antel, J., König, F. B., Brück, W. & Kuhlmann, T. Remyelination capacity of the MS brain decreases with disease chronicity. *Neurology* **72**, 1914–1921 (2009).
79. Immunopathology of multiple sclerosis: Report on an international meeting held at the Institute of Neurology of the University of Vienna. *J. Neuroimmunol.* **86**, 213–217 (1998).
80. Prineas, J. W. *et al.* Multiple sclerosis. Pathology of recurrent lesions. *Brain* **116 (Pt 3)**, 681–693 (1993).
81. Thompson, A. J. *et al.* Major differences in the dynamics of primary and secondary progressive multiple sclerosis. *Ann. Neurol.* **29**, 53–62 (1991).
82. Albert, M., Antel, J., Brück, W. & Stadelmann, C. Extensive cortical remyelination in patients with chronic multiple sclerosis. *Brain Pathol.* **17**, 129–138 (2007).
83. Patrikios, P. *et al.* Remyelination is extensive in a subset of multiple sclerosis patients. *Brain* **129**, 3165–3172 (2006).
84. Bieber, A. J., Kerr, S. & Rodriguez, M. Efficient central nervous system remyelination requires T cells. *Ann. Neurol.* **53**, 680–684 (2003).
85. Mason, J. L., Suzuki, K., Chaplin, D. D. & Matsushima, G. K. Interleukin-1beta promotes repair of the CNS. *J. Neurosci.* **21**, 7046–7052 (2001).

86. Chari, D. M., Zhao, C., Kotter, M. R., Blakemore, W. F. & Franklin, R. J. M. Corticosteroids delay remyelination of experimental demyelination in the rodent central nervous system. *J. Neurosci. Res.* **83**, 594–605 (2006).
87. Arnett, H. A. *et al.* TNF alpha promotes proliferation of oligodendrocyte progenitors and remyelination. *Nat. Neurosci.* **4**, 1116–1122 (2001).
88. Setzu, A. *et al.* Inflammation stimulates myelination by transplanted oligodendrocyte precursor cells. *Glia* **54**, 297–303 (2006).
89. Kotter, M. R., Zhao, C., van Rooijen, N. & Franklin, R. J. M. Macrophage-depletion induced impairment of experimental CNS remyelination is associated with a reduced oligodendrocyte progenitor cell response and altered growth factor expression. *Neurobiol. Dis.* **18**, 166–175 (2005).
90. Franklin, R. J. M. & Goldman, S. A. Glia Disease and Repair-Remyelination. *Cold Spring Harb. Perspect. Biol.* **7**, a020594 (2015).
91. Chang, A., Nishiyama, A., Peterson, J., Prineas, J. & Trapp, B. D. NG2-positive oligodendrocyte progenitor cells in adult human brain and multiple sclerosis lesions. *J. Neurosci.* **20**, 6404–6412 (2000).
92. Chen, C. Z., Neumann, B., Förster, S. & Franklin, R. J. M. Schwann cell remyelination of the central nervous system: why does it happen and what are the benefits? *Open Biol.* **11**, 200352 (2021).
93. Ghezzi, L. *et al.* Schwann Cell Remyelination in the Multiple Sclerosis Central Nervous System. *Lab. Invest.* **103**, 100128 (2023).
94. Patani, R., Balaratnam, M., Vora, A. & Reynolds, R. Remyelination can be extensive in multiple sclerosis despite a long disease course. *Neuropathol. Appl. Neurobiol.* **33**, 277–287 (2007).
95. Bramow, S. *et al.* Demyelination versus remyelination in progressive multiple sclerosis. *Brain* **133**, 2983–2998 (2010).

96. Prineas, J. W., Barnard, R. O., Kwon, E. E., Sharer, L. R. & Cho, E. S. Multiple sclerosis: remyelination of nascent lesions. *Ann. Neurol.* **33**, 137–151 (1993).

97. Mews, I., Bergmann, M., Bunkowski, S., Gullotta, F. & Brück, W. Oligodendrocyte and axon pathology in clinically silent multiple sclerosis lesions. *Mult. Scler.* **4**, 55–62 (1998).

98. Hametner, S. *et al.* Iron and neurodegeneration in the multiple sclerosis brain. *Ann. Neurol.* **74**, 848–861 (2013).

99. Holley, J. E., Gveric, D., Newcombe, J., Cuzner, M. L. & Gutowski, N. J. Astrocyte characterization in the multiple sclerosis glial scar. *Neuropathol. Appl. Neurobiol.* **29**, 434–444 (2003).

100. Ridet, J. L., Malhotra, S. K., Privat, A. & Gage, F. H. Reactive astrocytes: cellular and molecular cues to biological function. *Trends Neurosci.* **20**, 570–577 (1997).

101. Piétu, G. *et al.* The Genexpress IMAGE knowledge base of the human brain transcriptome: a prototype integrated resource for functional and computational genomics. *Genome Res.* **9**, 195–209 (1999).

102. Williams, C. G., Lee, H. J., Asatsuma, T., Vento-Tormo, R. & Haque, A. An introduction to spatial transcriptomics for biomedical research. *Genome Med.* **14**, 68 (2022).

103. Adams, M. D. *et al.* Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651–1656 (1991).

104. Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. Serial analysis of gene expression. *Science* **270**, 484–487 (1995).

105. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470 (1995).

106. McGettigan, P. A. Transcriptomics in the RNA-seq era. *Curr. Opin. Chem. Biol.* **17**, 4–11 (2013).

107. Bainbridge, M. N. *et al.* Analysis of the prostate cancer cell line LNCaP transcriptome using a sequencing-by-synthesis approach. *BMC Genomics* **7**, 246 (2006).

108. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
109. Sultan, M. *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**, 956–960 (2008).
110. Wilhelm, B. T. *et al.* Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**, 1239–1243 (2008).
111. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
112. Tian, L., Chen, F. & Macosko, E. Z. The expanding vistas of spatial transcriptomics. *Nat. Biotechnol.* **41**, 773–782 (2023).
113. Schirmer, L. *et al.* Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75–82 (2019).
114. Absinta, M. *et al.* A lymphocyte-microglia-astrocyte axis in chronic active multiple sclerosis. *Nature* **597**, 709–714 (2021).
115. Haque, A., Engel, J., Teichmann, S. A. & Lönnberg, T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med.* **9**, 75 (2017).
116. Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nat. Rev. Genet.* **24**, 550–572 (2023).
117. Lacar, B. *et al.* Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* **7**, 11022 (2016).
118. Aldridge, S. & Teichmann, S. A. Single cell transcriptomics comes of age. *Nat. Commun.* **11**, 4307 (2020).
119. Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599–604 (2018).
120. Mereu, E. *et al.* Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat. Biotechnol.* **38**, 747–755 (2020).

121. Baek, S. & Lee, I. Single-cell ATAC sequencing analysis: From data preprocessing to hypothesis generation. *Comput. Struct. Biotechnol. J.* **18**, 1429–1439 (2020).
122. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
123. Han, A., Glanville, J., Hansmann, L. & Davis, M. M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* **32**, 684–692 (2014).
124. Larsson, L., Frisén, J. & Lundeberg, J. Spatially resolved transcriptomics adds a new dimension to genomics. *Nat. Methods* **18**, 15–18 (2021).
125. Xia, B. & Yanai, I. A periodic table of cell types. *Development* **146**, (2019).
126. Mincarelli, L., Lister, A., Lipscombe, J. & Macaulay, I. C. Defining Cell Identity with Single-Cell Omics. *Proteomics* **18**, e1700312 (2018).
127. Rao, A., Barkley, D., França, G. S. & Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature* **596**, 211–220 (2021).
128. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
129. Merritt, C. R. *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat. Biotechnol.* **38**, 586–599 (2020).
130. He, S. *et al.* High-plex multiomic analysis in FFPE at subcellular level by spatial molecular imaging. *bioRxiv* (2021) doi:10.1101/2021.11.03.467020.
131. Moffitt, J. R. *et al.* Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, (2018).
132. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598**, 137–143 (2021).
133. Chen, W.-T. *et al.* Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. *Cell* **182**, 976–991.e19 (2020).

134. Hunter, M. V., Moncada, R., Weiss, J. M., Yanai, I. & White, R. M. Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface. *Nat. Commun.* **12**, 6278 (2021).
135. Ke, R. *et al.* In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 857–860 (2013).
136. Jain, S., Dandy, D. S., Geiss, B. J. & Henry, C. S. Padlock probe-based rolling circle amplification lateral flow assay for point-of-need nucleic acid detection. *Analyst* **146**, 4340–4347 (2021).
137. Alon, S. *et al.* Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems. *Science* **371**, (2021).
138. Gyllborg, D. *et al.* Hybridization-based in situ sequencing (HybISS) for spatially resolved transcriptomics in human and mouse brain tissue. *Nucleic Acids Res.* **48**, e112 (2020).
139. Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nat. Methods* **19**, 534–546 (2022).
140. Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. Visualization of single RNA transcripts in situ. *Science* **280**, 585–590 (1998).
141. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell in situ RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014).
142. Wang, G., Moffitt, J. R. & Zhuang, X. Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. *Sci. Rep.* **8**, 4847 (2018).
143. Beliveau, B. J. *et al.* Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 21301–21306 (2012).
144. Huang, Z.-L. *et al.* Localization-based super-resolution microscopy with an sCMOS camera. *Opt. Express* **19**, 19156–19168 (2011).
145. Saurabh, S., Maji, S. & Bruchez, M. P. Evaluation of sCMOS cameras for detection and localization of single Cy5 molecules. *Opt. Express* **20**, 7338–7349 (2012).
146. Website. https://www.10xgenomics.com/ spatial-transcriptomics/.

147. Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).

148. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

149. Stickels, R. R. *et al.* Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* **39**, 313–319 (2021).

150. Vickovic, S. *et al.* High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* **16**, 987–990 (2019).

151. Liu, Y. *et al.* High-Spatial-Resolution Multi-Omics Sequencing via Deterministic Barcoding in Tissue. *Cell* **183**, 1665–1681.e18 (2020).

152. Chen, A. *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792.e21 (2022).

153. Cho, C.-S. *et al.* Microscopic examination of spatial transcriptome using Seq-Scope. *Cell* **184**, 3559–3572.e22 (2021).

154. Fu, X. *et al.* Continuous polony gels for tissue mapping with high resolution and RNA capture efficiency. *bioRxiv* (2021) doi:10.1101/2021.03.17.435795.

155. Biancalani, T. *et al.* Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat. Methods* **18**, 1352–1362 (2021).

156. Cable, D. M. *et al.* Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, 517–526 (2022).

157. Kleshchevnikov, V. *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).

158. Saul, D. *et al.* A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. *Nat. Commun.* **13**, 4827 (2022).

159. Lerma-Martin, C. *et al.* Cell type mapping reveals tissue niches and interactions in subcortical multiple sclerosis lesions. *Nature Neuroscience* 1–12 (2024) doi:10.1038/s41593-024-01796-z.

160. Frischer, J. M. *et al.* Clinical and pathological insights into the dynamic nature of the white matter multiple sclerosis plaque. *Ann. Neurol.* **78**, 710–721 (2015).

161. Emery, B. Regulation of oligodendrocyte differentiation and myelination. *Science* **330**, 779–782 (2010).

162. Smith, B. C. *et al.* Astrocyte interferon-gamma signaling dampens inflammation during chronic central nervous system autoimmunity via PD-L1. *J. Neuroinflammation* **20**, 234 (2023).

163. Kann, O., Almouhanna, F. & Chausse, B. Interferon γ: a master cytokine in microglia-mediated neural network dysfunction and neurodegeneration. *Trends Neurosci.* **45**, 913–927 (2022).

164. Machado-Santos, J. *et al.* The compartmentalized inflammatory response in the multiple sclerosis brain is composed of tissue-resident CD8+ T lymphocytes and B cells. *Brain* **141**, 2066–2082 (2018).

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

166. Gautam, J., Miner, J. H. & Yao, Y. Loss of endothelial laminin α5 exacerbates hemorrhagic brain injury. *Transl. Stroke Res.* **10**, 705 (2019).

167. Zemskov, E. A. *et al.* NF-κB-dependent repression of Sox18 transcription factor requires the epigenetic regulators histone deacetylases 1 and 2 in acute lung injury. *Front. Physiol.* **13**, (2022).

168. Jiang, L., Zhu, X., Yang, H., Chen, T. & Lv, K. Bioinformatics Analysis Discovers Microtubular Tubulin Beta 6 Class V (TUBB6) as a Potential Therapeutic Target in Glioblastoma. *Front. Genet.* **11**, 566579 (2020).

169. Huang, T. *et al.* The physiological role of Motin family and its dysregulation in tumorigenesis. *J. Transl. Med.* **16**, 98 (2018).

170. Habeeb, O. *et al.* Epithelial membrane protein-2 expression is an early predictor of endometrial cancer development. *Cancer* **116**, 4718–4726 (2010).

171. Russo, E. *et al.* CD112 Regulates Angiogenesis and T Cell Entry into the Spleen. *Cells* **10**, (2021).

172. Zhao, J. *et al.* Identification of target genes in neuroinflammation and neurodegeneration after traumatic brain injury in rats. *PeerJ* **7**, e8324 (2019).

173. Macnair, W. *et al.* Single nuclei RNAseq stratifies multiple sclerosis patients into three distinct white matter glia responses. *bioRxiv* (2022) doi:10.1101/2022.04.06.487263.

174. Liu, H., Hu, Q., D'ercole, A. J. & Ye, P. Histone deacetylase 11 regulates oligodendrocyte-specific gene expression and cell development in OL-1 oligodendroglia cells. *Glia* **57**, 1–12 (2009).

175. Brockschnieder, D., Sabanay, H., Riethmacher, D. & Peles, E. Ermin, a myelinating oligodendrocyte-specific protein that regulates cell morphology. *J. Neurosci.* **26**, 757–762 (2006).

176. Miyazaki, Y. *et al.* Oligodendrocyte-derived LGI3 and its receptor ADAM23 organize juxtaparanodal Kv1 channel clustering for short-term synaptic plasticity. *Cell Rep.* **43**, 113634 (2024).

177. Cho, C. H. *et al.* ANKS1B encoded AIDA-1 regulates social behaviors by controlling oligodendrocyte function. *Nat. Commun.* **14**, 8499 (2023).

178. Linnerbauer, M., Wheeler, M. A. & Quintana, F. J. Astrocyte Crosstalk in CNS Inflammation. *Neuron* **108**, 608–622 (2020).

179. Schirmer, L., Schafer, D. P., Bartels, T., Rowitch, D. H. & Calabresi, P. A. Diversity and Function of Glial Cell Types in Multiple Sclerosis. *Trends Immunol.* **42**, 228–247 (2021).

180. Zhang, Q. *et al.* Single-cell analysis reveals dynamic changes of neural cells in developing human spinal cord. *EMBO Rep.* **22**, e52728 (2021).

181. Jacquet, B. V. *et al.* FoxJ1-dependent gene expression is required for differentiation of radial glia into ependymal cells and a subset of astrocytes in the postnatal brain. *Development* **136**, 4021–4031 (2009).

182. Berg, J. S. *et al.* Next generation massively parallel sequencing of targeted exomes to identify genetic mutations in primary ciliary dyskinesia: implications for application to clinical testing. *Genet. Med.* **13**, 218–229 (2011).

183. McKenzie, C. W. *et al.* CFAP54 is required for proper ciliary motility and assembly of the central pair apparatus in mice. *Mol. Biol. Cell* **26**, 3140–3149 (2015).

184. Teves, M. E. *et al.* Sperm-associated antigen-17 gene is essential for motile cilia function and neonatal survival. *Am. J. Respir. Cell Mol. Biol.* **48**, 765–772 (2013).

185. Huang, W., Hickson, L. J., Eirin, A., Kirkland, J. L. & Lerman, L. O. Cellular senescence: the good, the bad and the unknown. *Nat Rev Nephrol* **18**, 611–627 (2022).

186. Wallmeier, J. *et al.* De Novo Mutations in FOXJ1 Result in a Motile Ciliopathy with Hydrocephalus and Randomization of Left/Right Body Asymmetry. *Am. J. Hum. Genet.* **105**, 1030–1039 (2019).

187. Dimitrov, D. *et al.* Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat. Commun.* **13**, 3224 (2022).

188. Armingol, E., Baghdassarian, H. M. & Lewis, N. E. The diversification of methods for studying cell-cell interactions and communication. *Nat. Rev. Genet.* **25**, 381–400 (2024).

189. Ren, W., Zhao, L., Sun, Y., Wang, X. & Shi, X. HMGB1 and Toll-like receptors: potential therapeutic targets in autoimmune diseases. *Mol. Med.* **29**, 117 (2023).

190. Hofmann, A. *et al.* Myeloid cell iron uptake pathways and paramagnetic rim formation in multiple sclerosis. *Acta Neuropathol.* **146**, 707–724 (2023).

191. Yang, H. *et al.* Identification of CD163 as an antiinflammatory receptor for HMGB1-haptoglobin complexes. *JCI Insight* **1**, (2016).

192. Humphries, J. D. & Humphries, M. J. CD14 is a ligand for the integrin alpha4beta1. *FEBS Lett.* **581**, 757–763 (2007).

193. Kukanja, P. *et al.* Cellular architecture of evolving neuroinflammatory lesions and multiple sclerosis pathology. *Cell* **187**, 1990–2009.e19 (2024).

194. Jäkel, S. *et al.* Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).

195. Kaufmann, M. *et al.* Identifying CNS-colonizing T cells as potential therapeutic targets to prevent progression of multiple sclerosis. *Med* **2**, 296–312.e8 (2021).

196. Kaufmann, M. *et al.* Identification of early neurodegenerative pathways in progressive multiple sclerosis. *Nat. Neurosci.* **25**, 944–955 (2022).

197. Kuppe, C. *et al.* Spatial multi-omic map of human myocardial infarction. *Nature* **608**, 766–777 (2022).

198. Parigi, S. M. *et al.* The spatial transcriptomic landscape of the healing mouse intestine following damage. *Nat. Commun.* **13**, 828 (2022).

199. Huuki-Myers, L. A. *et al.* A data-driven single-cell and spatial transcriptomic map of the human prefrontal cortex. *Science* **384**, eadh1938 (2024).

200. Shi, Y. *et al.* Endothelium-targeted overexpression of heat shock protein 27 ameliorates blood-brain barrier disruption after ischemic brain injury. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1243–E1252 (2017).

201. Ramaglia, V., Florescu, A., Zuo, M., Sheikh-Mohamed, S. & Gommerman, J. L. Stromal Cell-Mediated Coordination of Immune Cell Recruitment, Retention, and Function in Brain-Adjacent Regions. *J. Immunol.* **206**, 282–291 (2021).

202. Engelhardt, B. *et al.* Vascular, glial, and lymphatic immune gateways of the central nervous system. *Acta Neuropathol.* **132**, 317–338 (2016).

203. Larochelle, C., Alvarez, J. I. & Prat, A. How do immune cells overcome the blood-brain barrier in multiple sclerosis? *FEBS Lett.* **585**, 3770–3780 (2011).

204. Fransen, N. L. *et al.* Tissue-resident memory T cells invade the brain parenchyma in multiple sclerosis white matter lesions. *Brain* **143**, 1714–1730 (2020).

205. Lalor, S. J. & Segal, B. M. Lymphoid chemokines in the CNS. *J. Neuroimmunol.* **224**, 56–61 (2010).

206. Mohan, H. *et al.* Extracellular matrix in multiple sclerosis lesions: Fibrillar collagens, biglycan and decorin are upregulated and associated with infiltrating immune cells. *Brain Pathol.* **20**, 966–975 (2010).

207. Bear, R. M. & Caspary, T. Uncovering cilia function in glial development. *Ann. Hum. Genet.* **88**, 27 (2024).

208. Sipos, É., Komoly, S. & Ács, P. Quantitative Comparison of Primary Cilia Marker Expression and Length in the Mouse Brain. *J. Mol. Neurosci.* **64**, 397–409 (2018).

209. Bishop, G. A., Berbari, N. F., Lewis, J. & Mykytyn, K. Type III adenylyl cyclase localizes to primary cilia throughout the adult mouse brain. *J. Comp. Neurol.* **505**, 562–571 (2007).

210. Sterpka, A. *et al.* Diverged morphology changes of astrocytic and neuronal primary cilia under reactive insults. *Mol. Brain* **13**, 28 (2020).

211. Wang, L. *et al.* The diversified astrocyte developmental programs are modulated by primary ciliary signaling. *bioRxiv* (2024) doi:10.1101/2024.03.17.585433.

212. Muhamad, N. A. *et al.* Astrocyte-specific inhibition of primary cilium functions improves cognitive impairment during neuroinflammation by suppressing A1 astrocyte differentiation. *bioRxiv* (2023) doi:10.1101/2023.10.01.560403.

213. Venkatesh, D. Primary cilia. *J. Oral Maxillofac. Pathol.* **21**, 8–10 (2017).

214. Mill, P., Christensen, S. T. & Pedersen, L. B. Primary cilia as dynamic and diverse signalling hubs in development and disease. *Nat. Rev. Genet.* **24**, 421–441 (2023).

215. Patir, A. *et al.* The transcriptional signature associated with human motile cilia. *Sci. Rep.* **10**, 10814 (2020).

216. Ning, K. *et al.* Cilia-associated wound repair mediated by IFT88 in retinal pigment epithelium. *Sci. Rep.* **13**, 8205 (2023).

217. Kopinke, D., Roberson, E. C. & Reiter, J. F. Ciliary Hedgehog Signaling Restricts Injury-Induced Adipogenesis. *Cell* **170**, 340–351.e12 (2017).

218. Hosio, M. *et al.* Primary Ciliary Signaling in the Skin-Contribution to Wound Healing and Scarring. *Front Cell Dev Biol* **8**, 578384 (2020).

219. Villalobos, E. *et al.* Fibroblast Primary Cilia Are Required for Cardiac Fibrosis. *Circulation* **139**, 2342–2357 (2019).

220. Chrystal, P. W. *et al.* The inner junction protein CFAP20 functions in motile and non-motile cilia and is critical for vision. *Nat. Commun.* **13**, 6595 (2022).

221. Bormann, D. *et al.* Single nucleus RNA sequencing reveals glial cell type-specific responses to ischemic stroke. *bioRxiv* (2023) doi:10.1101/2023.12.26.573302.

222. Weber, C. E. *et al.* Long-term dynamics of multiple sclerosis iron rim lesions. *Mult. Scler. Relat. Disord.* **57**, 103340 (2022).

223. Wang, M. *et al.* MEIS2 regulates endothelial to hematopoietic transition of human embryonic stem cells by targeting TAL1. *Stem Cell Res. Ther.* **9**, 340 (2018).

224. Li, Y. *et al.* CCT5 induces epithelial-mesenchymal transition to promote gastric cancer lymph node metastasis by activating the Wnt/β-catenin signalling pathway. *Br. J. Cancer* **126**, 1684–1694 (2022).

225. Iser, I. C., Lenz, G. & Wink, M. R. EMT-like process in glioblastomas and reactive astrocytes. *Neurochem. Int.* **122**, 139–143 (2019).

226. Injury-specific signals elicit plasticity in glial cells from patients with brain injury. *Nat. Med.* **29**, 3018–3019 (2023).

227. Etzerodt, A. & Moestrup, S. K. CD163 and inflammation: biological, diagnostic, and therapeutic aspects. *Antioxid. Redox Signal.* **18**, 2352–2363 (2013).

228. Paudel, Y. N. *et al.* HMGB1: A Common Biomarker and Potential Target for TBI, Neuroinflammation, Epilepsy, and Cognitive Dysfunction. *Front. Neurosci.* **12**, 628 (2018).

229. Ved, R. *et al.* Disulfide HMGB1 acts via TLR2/4 receptors to reduce the numbers of oligodendrocyte progenitor cells after traumatic injury in vitro. *Sci. Rep.* **11**, 6181 (2021).

230. Lin, H. *et al.* Delineation of colorectal cancer ligand-receptor interactions and their roles in the tumor microenvironment and prognosis. *J. Transl. Med.* **19**, 497 (2021).
231. Tabata, H. *et al.* Erratic and blood vessel-guided migration of astrocyte progenitors in the cerebral cortex. *Nat. Commun.* **13**, 6571 (2022).
232. Yamamoto, H. *et al.* Integrin β1 controls VE-cadherin localization and blood vessel stability. *Nat. Commun.* **6**, 6429 (2015).
233. Hakanpaa, L. *et al.* Endothelial destabilization by angiopoietin-2 via integrin β1 activation. *Nat. Commun.* **6**, 1–12 (2015).
234. Nurminskaya, M. V. & Belkin, A. M. Cellular functions of tissue transglutaminase. *Int. Rev. Cell Mol. Biol.* **294**, 1–97 (2012).
235. Kawecki, C., Lenting, P. J. & Denis, C. V. von Willebrand factor and inflammation. *J. Thromb. Haemost.* **15**, 1285–1294 (2017).
236. Glatigny, S., Wagner, C. A. & Bettelli, E. Cutting Edge: Integrin α4 Is Required for Regulatory B Cell Control of Experimental Autoimmune Encephalomyelitis. *J. Immunol.* **196**, 3542–3546 (2016).
237. Glatigny, S., Duhen, R., Arbelaez, C., Kumari, S. & Bettelli, E. Integrin alpha L controls the homing of regulatory T cells during CNS autoimmunity in the absence of integrin alpha 4. *Sci. Rep.* **5**, 7834 (2015).
238. Yednock, T. A. *et al.* Prevention of experimental autoimmune encephalomyelitis by antibodies against alpha 4 beta 1 integrin. *Nature* **356**, 63–66 (1992).
239. Pietra, G., Romagnani, C., Manzini, C., Moretta, L. & Mingari, M. C. The emerging role of HLA-E-restricted CD8+ T lymphocytes in the adaptive immune response to pathogens and tumors. *J. Biomed. Biotechnol.* **2010**, 907092 (2010).