Studying Myeloid Cell Heterogeneity After Spinal Cord Injury via Time-Resolved Single-Cell RNA Sequencing

Regan N Hamel

St Edmund's College

Department of Clinical Neurosciences

University of Cambridge

August 21, 2021

This dissertation is submitted for the degree of Doctor of Philosophy.

DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. Additionally, to specify when I performed the work, I use the pronoun "I", whereas when work was performed through collaboration, I use the pronoun "we".

I confirm that this dissertation is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

The length of this dissertation does not exceed 60,000 words, excluding tables, footnotes, references, and appendices and has fewer than 150 figures as required by the Degree Committee for Medical Sciences.

Regan Hamel

STUDYING MYELOID CELL HETEROGENEITY AFTER SPINAL CORD INJURY VIA TIME-RESOLVED SINGLE-CELL RNA SEQUENCING

ABSTRACT

Spinal cord injury (SCI) is a devastating pathology that affects thousands of individuals annually, resulting in the requirement for long-term physical and medical care and thus significant personal, societal, and economic burdens. The SCI pathology is characterised by an initial mechanical insult, followed by a spatiotemporally dynamic secondary injury. Decades of research have worked to assemble a general picture of this secondary pathology. We now understand that compared to the normal wound healing observed in the periphery, tissue recovery after SCI is dysregulated and results in a chronic wound state characterized by persistent inflammation and functional deficits. The primary drivers of this inflammation are central nervous system (CNS) resident microglia and infiltrating myeloid cells. However, the precise role of these myeloid cell subsets remains unclear as upon crossing the blood-spinal cord barrier (BSCB), infiltrating monocyte-derived macrophages may take on the morphology of microglia, and upregulate canonical microglia markers, making the two populations difficult to distinguish.

In this PhD project, I employed single-cell RNA sequencing (scRNAseq) to deconvolute the complex heterogeneity of infiltrating and resident myeloid cells in mouse models of thoracic contusion SCI at an unprecedented resolution. To fully appreciate the temporal dynamics of the pathology, I collected samples across the acute, subacute, and early chronic phases of SCI, plus a sham-injured control. Recent experiments have demonstrated that CNS infiltrating macrophages also take on the transcriptional profiles of microglia, which led me to question whether I had accurately annotated infiltrating macrophages in the dataset. To address this, I repeated the experiment with a transgenic fate-mapping mouse line then integrated these two datasets to generate a time-resolved SCI myeloid cell atlas with definitive ontogeny labelling. With this dataset I generated a putative time resolved map of myeloid cell dynamics across the SCI pathology. Through collaboration, I was also able to verify the expression of select genes via single-molecule fluorescent *in situ* hybridization (smFISH) and immunofluorescence (IF). A key observation was the persistence of a pro-inflammatory foam cell-like state in both microglia and macrophages, which may contribute to the non-resolving chronic injury. Future studies might investigate the functional relevance of this population, and its suitability as a therapeutic target to reduce the long-term disabilities of SCI patients.

REGAN N HAMEL

ACKNOWLEDGEMENTS

Thanks to Stefano Pluchino for his supervision. Bryan, Veronica, and Yutong, thank you all for making me a better teacher. Aletta, Greg, and Bristena, cheers for the inspiration, your amazing support, and your friendship. Jayden, thanks for the jokes. Thanks to all the other members of the Pluchino lab for their contributions over the years. Thank you Aviva for casual, but ever insightful conversations in the kitchen and the dissection room, and for being a wonderful example of a strong woman in science. Thanks to John Marioni for his incredible support as a secondary supervisor and for being a model of considerate and effective leadership. Karsten, Shila, Leah, Andrian and the rest of the Marioni group, I truly appreciated all your help and advice and your manifestation of the type of positive work environment I aspire to contribute to.

A quick note of appreciation to Kasia Katarzyna, who was always a smiling face I looked forward to seeing at the end of a long day of single-cell isolation.

To our collaborators, especially Stefano Stifani, Louise Thiry, and Katherine Ridley, a special thanks for their insight and encouragement.

Thanks to the Cambridge Trust, Rosetrees Trust, Wings For Life, and Canada Scholarship Trust for sponsoring me throughout this journey.

Cheers to Michelle, Liz, Amar, and Ana who kept me close to North America, and Carrie, Priyanka, Tom, Emma, Dhruv, and Gloria and all the lovely faces in Cambridge who make it feel like home.

Of course, I must express my deep appreciation for all my loving Canadian, American, and Swiss families who played a crucial role in my academic success, including of course my Mamma, for telling me to Reganize it, Pops, for making me strong and independent, and Brenna, for being the most inspirational sisterfriend anyone could ever ask for. Finally, thank you Clay, for your unwavering support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Acronyms

AD	Alzheimer's disease
ALS	amyotrophic lateral sclerosis
ATAC	Assay for Transposase-Accessible Chromatin
BBB	blood-brain barrier
BMDM	bone marrow derived macrophages
Bp	base pair
BSCB	blood-spinal cord barrier
CNS	central nervous system
CSPGs	chondroitin sulphate proteoglycans
CST	corticospinal tract
Ctrl	control
CyTOF	cytometry by time of flight
DAM	disease-associated microglia (Keren-Shaul et al., 2017)
DAMPs	damage-associated molecular patterns
DEGs	differentially expressed genes
Dpi	day post SCI
EAE	experimental autoimmune encephalomyelitis
FACS	fluorescent activated cell sorting
FFPE	formalin-fixed paraffin-embedded
GEM	gel bead in emulsion
GO	gene ontology
HSC	haematopoietic stem cells
HVGs	highly variable genes
IF	immunofluorescence
IHC	immunohistochemistry
INTACT	isolation of nuclei tagged in specific cell types
KO	knock out
log_2FC	log_2 fold change
MGnD	disease-associated microglia (Krasemann et al., 2017)
MHC	major histocompatibility complex
MNN	mutual nearest neighbours
MS	multiple sclerosis
NASC	normal appearing spinal cord
NGS	next generation sequencing
NPCs	neural progenitor cells
OXPHOS	oxidative phosphorylation
PC	principal component
PCA	principal component analysis
PFA	paraformaldehyde
QC	quality control
RFP	red fluorescent protein/tdTomato
ROI	region of interest
scRNAseq	single-cell RNA sequencing
smFISH	single-molecule fluorescent *in situ* hybridization
snRNAseq	single-nuclear RNA sequencing

SNN	shared nearest neighbour
t-SNE	t-stochastic neighbour embedding
UMAP	Uniform manifold approximation and projection
UMI	unique molecular identifier
WT	wild-type
WTA	Whole Transcriptome Atlas
YFP	yellow fluorescent protein

Chapter 1 INTRODUCTION

"This job is a great scientific adventure. But it's also a great human adventure. Mankind has made giant steps forward. However, what we know is really very, very little compared to what we still have to know."

— Fabiola Gianotti

1.1 Spinal Cord Injury

1.1.1 Epidemiology

Spinal cord injury (SCI) is a highly debilitating pathology affecting thousands of individuals with regional annual incidences ranging from 11-16 per 100,000 (James et al., 2019; Singh et al., 2014). Traumatic SCI, most frequently caused by motor vehicle accidents and falls, is the most common form of SCI, and is typically more prevalent in men than women, with a minimum ratio of 2:1. However, this pathology can also be inflicted by infarction, genetic defects, or cancer (DeVivo, 2012). SCI primarily occurs in adults over 70 or in young adults (15-30 years) with no previous health conditions. Of those afflicted, up to 80% will survive to the chronic stage of the pathology, with younger patients having substantially better outcomes (Alizadeh et al., 2019). Irrespective of age, patients typically experience a significant reduction in both their quality of life and life expectancy (Ahuja et al., 2017; Singh et al., 2014).

The symptoms of SCI are dependent upon the anatomical location and severity of the lesion. The most common location of SCI is cervical, with nearly 50% of cases occurring at this level (Alizadeh et al., 2019). Thoracic and lumbar injuries are not uncommon, at 35% and 11% of cases, respectively. General symptoms include the reduction or complete deficit of sensory and motor functions below the level of injury (Alizadeh et al., 2019; Hou & Rabchevsky, 2014). Autonomic functions, such as respiration, urination, gastrointestinal motility, and thermoregulation, can also be affected. Thus, the severe and enduring disabilities caused by SCI, and the requirement for long-term health care, are not only detrimental to the patient's standard of living, but also come with significant personal, societal, and economic burdens.

Despite the evident impact, few therapeutic options are available for the treatment of central nervous system (CNS) pathologies, including SCI. Regenerative therapies, such as stem cell transplantation, represent promising avenues for future CNS therapeutics (Smith et al., 2020). However, we are currently limited by our poor understanding of the injury microenvironment. Non-neuronal cells are major players in this, and optimal strategies to address SCI will likely involve the therapeutic modulation of the injury environment at a cellular and molecular level. A time-resolved understanding of these heterogenous cell types and states contributing to the SCI pathology is essential.

1.1.2 An Overview of the Human Primary and Secondary Injuries

Most of our understanding of human SCI arises from mechanistic studies in animal models of SCI. Links between these experiments and the human condition are then inferred based on histopathology from post-mortem tissue, or minimally invasive techniques, such as cerebral spinal fluid sampling or MRI.

The SCI pathology can be described in two dynamic stages: primary and secondary. The primary injury is the initial mechanical insult that compromises the structural integrity of the vertebrae, blood-spinal cord barrier (BSCB), and spinal cord. This event typically damages the spinal cord parenchyma, but only very rarely does it completely sever the tissue (Alizadeh et al., 2019). There are four main types of SCI, listed in order of decreasing frequency: contusion with compression, contusion and transient compression, distraction, and laceration or transection (Alizadeh et al., 2019). Each of these injuries result in the immediate disruption of the microvasculature, vasospasm, systemic hypotension, necrosis, axonal damage, ionic imbalance, and the extracellular accumulation of neurotransmitters leading to excitotoxity (Alizadeh et al., 2019; Rowland et al., 2008; Silva et al., 2014).

The primary injury sparks the secondary injury, a cascade of biological events that spread the injury into the adjacent tissue. The secondary injury is proportional to the severity of the primary injury and can be divided into three stages: acute, subacute, and chronic (Figure 1.1). In humans, these stages are ≤ 48 hours, ≤ 2 weeks, and > 2 weeks, respectively (Tyler et al., 2013; Witiw & Fehlings, 2015) (Figure 1.2). The acute phase comprises the continuation of the immediate-onset processes of the primary injury plus: ischaemia; calcium influx leading to calpain-mediated protein degradation, mitochondrial failure-induced oxidative damage, and apoptosis; free radical formation causing membrane lipid peroxidation and a chain reaction that oxidises all the unsaturated lipids in the area, ultimately leading to respiratory and metabolic failure-induced cell death; myelin damage; oedema; and neuroinflammation (Alizadeh et al., 2019; Kopper & Gensel, 2018). Neuroinflammation begins with the release of damage-associated molecular patterns (DAMPs), namely alarmins, from dead and dying cells (Gadani et al., 2015). Alarmins signal to the undamaged cells in the perilesional area, consequently spreading the inflammatory response. Microglia and reactive astrocytes,

activated and attracted by the alarmins, trigger the innate immune response through cytokine production, resulting in the attraction of blood-borne neutrophils and monocytes, which enter the spinal cord through the disrupted BSCB and migrate to the lesion as early as 1 day post-injury (dpi) (Alizadeh et al., 2019; Milich et al., 2019). By 3 dpi, the inflammatory response also includes the recruitment and activation of T- and B-lymphocytes, a response which peaks around a month post-SCI. Thus, while the CNS cell population is drastically reduced through acute and subacute necrosis, Fas-mediated apoptosis, and lysosomal damage-induced necroptosis, the pro-inflammatory response works to recover cell numbers as early as 3 dpi by recruiting substantial volumes of peripheral* cells and inducing the proliferation, spatial reorganisation, and hypertrophy of glial cells, a process known as gliosis (Burda & Sofroniew, 2014).

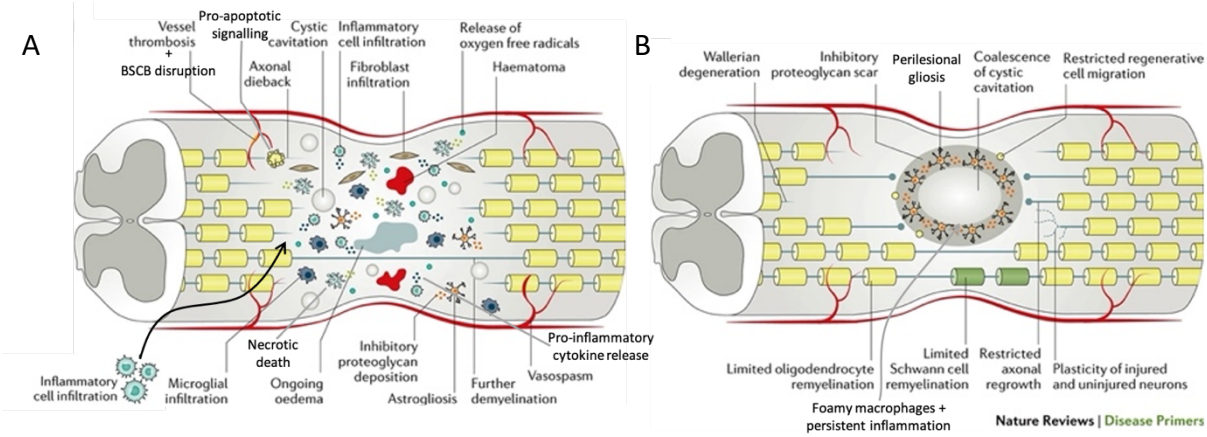


Figure 1.1 The secondary response to SCI. A) The subacute phase is complex and dynamic. B) The chronic phase is characterised by the maturation of the glial scar, limited regeneration, and the presence of pro-inflammatory foamy macrophages. Figure adapted from (Ahuja et al., 2017)

In the early subacute phase, blood-borne monocytes differentiate into dendritic cells and macrophages in the spinal cord parenchyma (Kigerl et al., 2009; Milich et al., 2019). Macrophages phagocytose debris and activate perivascular pericytes and fibroblasts, beginning the process of fibrosis (Zhu et al., 2015). These cells migrate to the lesion epicentre and form a fibrotic core, characterised by an extracellular matrix high in fibronectin, collagen, and laminin as well as anti-regenerative molecules such as semaphorins (Alizadeh et al., 2019). By 7 dpi, this fibrotic core begins to be surrounded by glial cells (Alizadeh et al., 2019; Sofroniew, 2015). This corralling, or the compacting of the lesion area and the formation of

* For the purposes of this dissertation, periphery refers to non-CNS

this protective barrier around the lesion, has been shown to limit the further spread of tissue damage in animal models (Bradbury & Burnside, 2019; Okada et al., 2018; Zhou et al., 2020). The glial-barrier, which ultimately becomes the glial scar, is a thick, dense, fibrillary structure comprising scar-forming astrocytes and microglia, NG2-expressing oligodendrocyte precursor cells, fibroblasts, ependymal cells, and pericytes (Bellver-Landete et al., 2019; Fleming et al., 2006; Liddelow & Barres, 2017). The ablation of the glial scar in the chronic phase of SCI in mouse models results in enlarged lesions and further tissue degeneration, suggesting that the cells within the fibrotic core remain active into the chronic phase (Anderson et al., 2016). Whether this also applies to the human condition has yet to be investigated. Like other aspects of SCI, most of our understanding of the glial scar and fibrotic core stems from animal models, namely rodents. The presence of both elements have been confirmed in humans, however there is substantial evidence that Schwann cells contribute to the human scar and fibrotic core in a way that is not observed in rodents (Orr & Gensel, 2018).

While important for limiting the spread of the SCI pathology, the glial scar also plays a negative role in regeneration. The scar itself can act as a physical barrier to axon growth, and its cellular components secrete inhibitory molecules, such as chondroitin sulphate proteoglycans (CSPGs), which limit axon regeneration, sprouting, and remyelination (Alizadeh et al., 2019; Tyler et al., 2013). Additionally, the glial scar does not resolve, persisting chronically in the spinal cord parenchyma (Cregg et al., 2014).

Beyond glial scarring, the subacute phase is also characterised by continued BSCB permeability, cell death, axonal retraction from the injury site, but also hypoxia-induced angiogenesis and revascularisation (Bradbury & Burnside, 2019; Whetstone et al., 2003). Subacute phase demyelination and Wallerian degeneration lead to an increased deposition of myelin debris and the generation of a lipid-dense microenvironment (X. Wang et al., 2015). This myelin debris is gradually cleared through phagocytosis, but this process takes years in the human SCI pathology and the debris has been associated with perpetuating autoimmunity and preventing oligodendrocyte differentiation, remyelination, and axonal regeneration (Kopper & Gensel, 2018). In peripheral nerve injury, myelin debris is resolved within weeks, suggesting this contributes to the dysregulation of the wound healing process observed in SCI (Gensel & Zhang, 2015).

The chronic phase of SCI is significantly less dynamic than the acute and subacute phases. The main event of this phase syringomyelia, or the formation of fluid-filled cystic cavities, at the lesion epicentre, void of glia and neurons and surrounded by white matter. These cavities expand over time, negatively impacts the SCI progression by preventing axonal regeneration and compressing intact axons (Biyani & Masry, 1994; Zhang et al., 2015). Beyond this, the chronic phase is characterised by the STAT3-mediated maturation of the glial scar (Escartin et al., 2019; Yang et al., 2020), myeloid and lymphoid-mediated inflammation (Beck et al., 2010), demyelination that extends to even a decade after the injury (Guest et al., 2005) but also spontaneous remyelination, and the continued secretion of CSPGs and subsequently minimal axonal sprouting (Buss et al., 2009). SCI is a non-resolving pathology; thus, these chronic features are endured indefinitely.

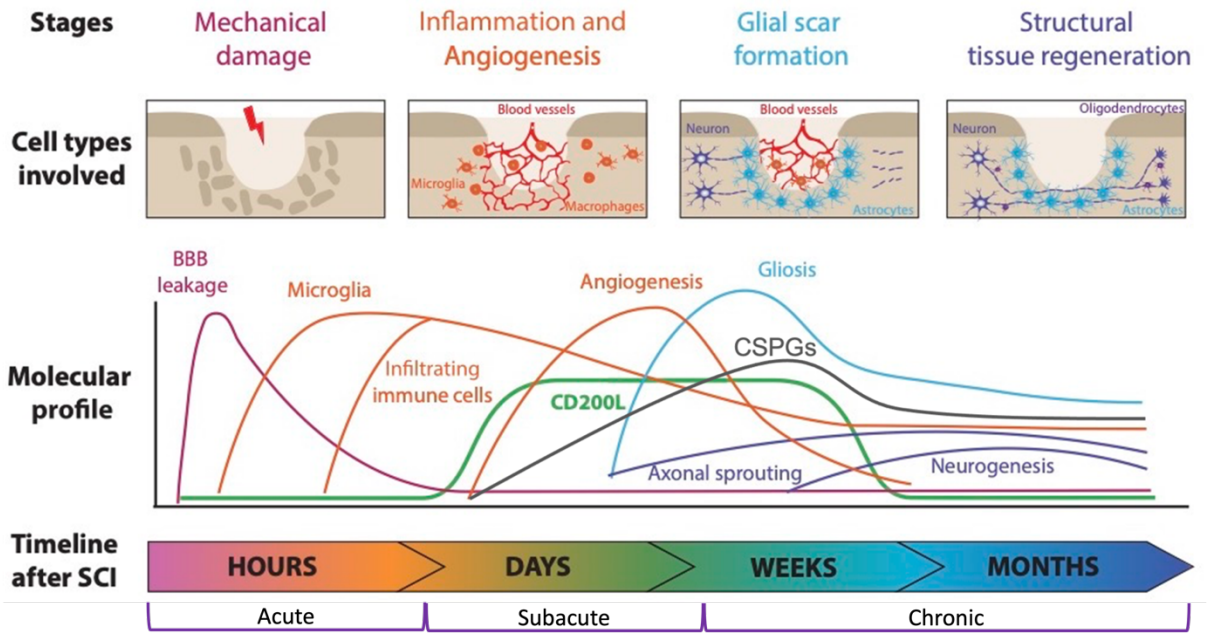


Figure 1.2 A summary of the temporally specific events that characterise the secondary injury. Figure adapted from (Rust & Kaiser, 2017).

1.1.3 A Closer Look at the Immune Response to SCI

The immune response to SCI is complex and dynamic with both beneficial and detrimental effects on tissue repair (Gaudet & Fonken, 2018). The response is mediated by the heterogenous activities of several cell types including astrocytes and *Cx3cr1*$^{+}$ leukocytes, namely lymphocytes, and myeloid cells (Hausmann, 2003).

ASTROCYTES

Under homeostatic conditions, astrocytes are the most abundant glial cell in the CNS and perform many diverse roles to maintain homeostasis (Bowes & Yip, 2014). Although astrocytes are not considered immune cells, they are immune-competent and in their activated state, upregulate major histocompatibility complex class II (MHCII), and thus are known modulators of both the innate and adaptive immune response (Bowes & Yip, 2014; Cekanaviciute & Buckwalter, 2016; Colombo & Farina, 2016; Farina et al., 2007; Okada et al., 2018). The exact immunomodulatory role of astrocytes after SCI is an area of active research, with animal models and *in vitro* studies outlining several signalling pathways that elicit either beneficial or detrimental responses (Colombo & Farina, 2016).

It is well-established that in the acute phase, astrocytes respond to environmental danger signals, such as alarmins, by migrating to the lesion epicentre and upregulating chemokine and cytokine production (Alizadeh et al., 2019). The astrocyte-modulated BSCB permeability and astrocyte-derived chemokines, MCP, CXCL1, CXCL2, and CCL2, are key drivers of the peripheral inflammatory myeloid cell infiltration (Cekanaviciute & Buckwalter, 2016). Astrocytes also influence the inflammatory responses of microglia and macrophages in the acute and subacute phases, through the secretion of both pro- and anti-inflammatory cytokines, subsequently inducing pro- and anti-inflammatory responses in these myeloid cells (Alizadeh et al., 2019; Yoshizaki et al., 2021). Furthermore, as previously discussed, astrocytes play a key role in limiting the spread of the inflammation by compacting the lesion and working with microglia to corral infiltrating inflammatory cells (Bellver-Landete et al., 2019; Zhou et al., 2020). The interruption of this process has been shown to limit functional recovery in rat models of contusion SCI (Okada et al., 2006). Finally, in the acute phase, reactive astrocytes also play a role in attracting lymphocytes to the lesion through the production of T- and B-lymphocyte chemokines, and the secretion of extravasation and homing molecules, such as VCAM (Alizadeh et al., 2019). However, further effects of astrocytes on the adaptive immune response are not yet fully understood in the context of human SCI (Cekanaviciute & Buckwalter, 2016).

LYMPHOCYTES

Lymphocytes, namely T- and B-cells, represent the adaptive immune response to SCI. The temporal dynamics of lymphocyte extravasation varies drastically between species. In humans, $CD8^+$ cytotoxic T-cells infiltrate the spinal cord as early as the acute phase, but drastically increase in quantity during the chronic phase (Fleming et al., 2006). $CD4^+$ helper T-cells follow the same dynamics, but in lower volumes. The presence of $CD20^+$ B-cells has not been observed in the injured human spinal cord. The temporal dynamics of lymphocytes in animal models of SCI will be discussed in *1.2 Animal Models*. The SCI pathology has been shown to provoke a chronic CNS autoimmune response, mediated by T- and B-cell responses to CNS antigens, especially from myelin-associated proteins (Alizadeh et al., 2019; Donnelly & Popovich, 2008). This autoimmune response further damages the spinal cord parenchyma through demyelination and the T-cell production of pro-inflammatory cytokines and chemokines that subsequently elicit pro-inflammatory responses in myeloid cells and facilitate the Fas-mediated apoptosis of neurons and oligodendrocytes (Alizadeh et al., 2019; Bowes & Yip, 2014). Impeding the T-cell response in animal models reduces tissue damage and improves functional outcomes (Alizadeh et al., 2019). As observed in a typical T-cell reaction, in the periphery, plasma cells differentiated from B-cells are also activated and secrete CNS autoantibodies, further exacerbating autoimmunity (Bowes & Yip, 2014; Donnelly & Popovich, 2008). The autoimmune response elicits further systemic effects in the periphery, but these will not be discussed here (Jones, 2014).

Unlike helper and cytotoxic T-cells, which are predominantly found in the chronic phase of SCI, recently, evidence of gamma delta T-cells in the cerebral spinal fluid of acute phase SCI patients has been reported alongside *in vivo* evidence of parenchymal infiltration in mouse models of SCI (Sun et al., 2018). These gamma delta T-cells are recruited by astrocyte- and microglia-secreted CCL2 and CCR2. These T-cells in turn secrete IFN-γ, a cytokine with both pro- and anti-inflammatory effects in the acute phase of SCI (Sun et al., 2018; Xu et al., 2021). However, in mouse models of SCI, gamma delta T-cells promote pro-inflammatory responses in macrophages, and the depletion of these lymphocytes improves functional recovery (Sun et al., 2018).

MYELOID CELLS

Myeloid cells, alongside natural killer cells and the complement system, are key mediators of the innate immune response to SCI (Alizadeh et al., 2019; Spiering, 2015). Understanding this response requires a knowledge of myeloid ontogeny. Myeloid is an umbrella term for a diverse set of immune cells that respond to tissue damage and pathogens (Figure 1.3) (Ransohoff & Cardona, 2010). Notably, not all myeloid cells share the same origins. During development, microglia arise from extra-embryonic yolk-sac blood islands and colonise the neuroepithelium where they maintain their population throughout adulthood via low-rate proliferation and apoptosis in the CNS parenchyma (Ginhoux et al., 2010; Masuda et al., 2020). These cells are generally long-lived, with a lifespan of several years in humans (Reu et al., 2017) and over a year in mice (Tay et al., 2017). Similarly, CNS-associated macrophages (CAMs), except for choroid plexus CAMs, have embryonic yolk-sac origins and long lifespans (Goldmann et al., 2016). Conversely, infiltrating neutrophils and monocyte-derived myeloid cells, such as macrophages and dendritic cells, arise from haematopoietic stem cells (HSC) in the bone marrow and spleen (Milich et al., 2019) and, under healthy conditions, are very short lived (1-3 days) (Kezic & McMenamin, 2008).

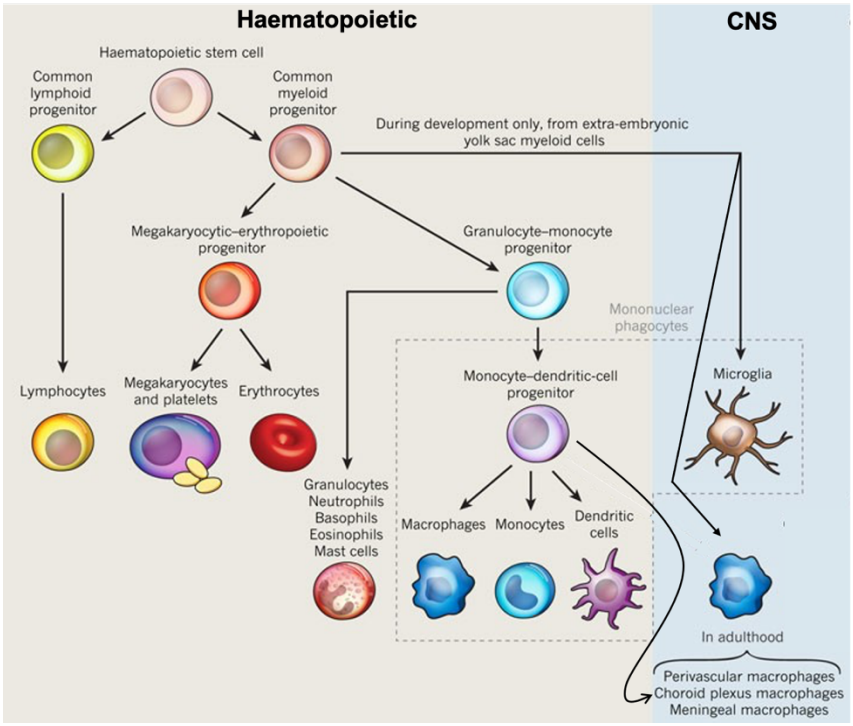


Figure 1.3 Myeloid cell ontogeny. Microglia and certain CAMs have a distinct ontogeny from the peripheral myeloid cells that infiltrate the CNS upon SCI. Figure adapted from (Ransohoff & Cardona, 2010).

NEUTROPHILS

Neutrophils are terminally differentiated, short-lived, blood-borne myeloid cells that are highly sensitive to chemokines, making them the typical first responders at the onset of a pathology (Zivkovic et al., 2021). Upon SCI, pattern recognition receptors (PRRs) on neutrophils rapidly detect and respond to DAMPs, resulting in their extravasation as early as 3 hours post-injury (Fleming et al., 2006; Zivkovic et al., 2021). By 1 dpi, neutrophils have fully infiltrated the spinal cord parenchyma and can be found throughout the white and grey matter. Thus, these cells represent the first infiltrating immune cells to arrive at the injury site (Zivkovic et al., 2021). By 5-10 dpi, neutrophils undergo apoptosis and are rarely observed in the tissue beyond the early subacute phase (Alizadeh et al., 2019; Fleming et al., 2006).

During this acute phase response, neutrophils are thought to perform three main functions, phagocytosis, degranulation, and the formation of extracellular traps, the details of which are not completely understood (Zivkovic et al., 2021). For example, it is well established that neutrophils actively phagocytose debris after SCI, but whether they aid in the clearance of regeneration-inhibiting myelin debris is unclear. Degranulation of neutrophils is a process through which neutrophils release their pathogen-eradicating contents, namely pro-inflammatory cytokines, serine proteases and free radicals, into the extracellular space (Alizadeh et al., 2019). It is unclear whether neutrophil degranulation results in further tissue damage, such as demyelination and further permeability of the BSCB (Zivkovic et al., 2021). Additionally, degranulation is assumed to occur after SCI based on evidence from other pathologies but has yet to be demonstrated. It has been shown that neutrophils do not contribute to post-SCI ROS production, which is evidence against degranulation (Bowes & Yip, 2014). Similarly, the formation of extracellular traps, networks of extracellular fibres mainly comprising DNA, designed to entrap pathogens, is assumed to occur but has yet to be described in SCI (Neirinckx et al., 2014; Zivkovic et al., 2021). Notably, despite the seemingly detrimental effects of neutrophils post-SCI, their ablation worsens functional outcomes in mouse models of SCI (Stirling et al., 2009), suggesting they play an important role setting the stage for the limited repair that does occur later in the pathology (Alizadeh et al., 2019).

MONOCYTE-DERIVED CELLS

Monocytes are $CD45^+/CD11b^+/CCR2^+$ blood-borne myeloid cells with the ability to differentiate into tissue macrophages and dendritic cells. As such, they are members of the innate immune system but also mediate adaptive immunity (Guilliams et al., 2014). In the acute phase of SCI, monocytes infiltrate the spinal cord parenchyma through the permeabilised BSCB (Zhou et al., 2014). $CD8^+$ monocytes have been reported to extravasate and migrate to the SCI site from 1-3 dpi (Fleming et al., 2006). By 3 dpi, monocytes begin to differentiate into dendritic cells and macrophages. Like neutrophils, little is known about the role of monocyte-derived dendritic cells in the SCI pathology. Dendritic cells are phagocytes and professional antigen-presenting cells via MHCII molecules and are thus known to play important roles in linking the innate and adaptive immune responses (Trivedi et al., 2006; Vaughn et al., 2013). They promote CD4+ T-cell proliferation and trigger an autoimmune response, but they also secrete neurotropic factors that improve neurogenesis (Bowes & Yip, 2014). In rodent and non-human primate models of SCI, enriching the lesion site with transplanted dendritic cells reduced demyelination and neuronal death, increased axonal sprouting, and improved functional outcomes (Yaguchi et al., 2009). However, the exact functions of endogenous dendritic cells after SCI, particularly in the human pathology, is not well understood.

Monocyte-derived macrophages populate the injured spinal cord, beginning at 3 dpi, as described, and peaking by 7 dpi (Milich et al., 2019). Macrophages remain present in the spinal cord well into the chronic phase, albeit at lower levels (Alizadeh et al., 2019; Fleming et al., 2006). It has been suggested that these acute phase macrophages arise from monocyte reservoirs in the spleen, while the chronic phase macrophages represent a second wave of bone marrow monocyte-derived macrophage infiltration beginning at 14 dpi and peaking around 60 dpi (Milich et al., 2019). However, this has yet to be investigated in the human condition. Furthermore, there is increasing evidence to support the self-renewal of macrophages after SCI, and it remains unclear how these two elements contribute to macrophage numbers after SCI (Greenhalgh & David, 2014; Milich et al., 2019).

Regardless of the precise details of their origin, monocyte-derived macrophages in the spinal cord parenchyma migrate to the lesion epicentre where they play important roles in

phagocytosing dying cells and myelin debris (Alizadeh et al., 2019; Bellver-Landete et al., 2019; Kong & Gao, 2017; X. Wang et al., 2015). Based on functional studies from animal models of SCI, macrophages that scavenge myelin debris are more likely to undergo apoptotic and necrotic cell death, compared to their microglia counterparts (Kong & Gao, 2017). In further support of ineffective myelin debris breakdown in macrophages upon SCI is the presence of foam cells, lipid-laden macrophages originally identified in atherosclerosis and described as pro-inflammatory and cytotoxic with reduced phagocytic capabilities. These cells have been observed in the grey matter of human SCI from 5 dpi to at least 1 year post-injury, suggesting they work to perpetuate the inflammatory response (Bradbury & Burnside, 2019; Fleming et al., 2006; X. Wang et al., 2015; Zhu et al., 2017). Notably, no distinction between activated microglia and macrophages was made in this histopathological analysis (Fleming et al., 2006). This inadequate breakdown of myelin debris likely contributes to the chronic adaptive immune response (Bradbury & Burnside, 2019; Kopper & Gensel, 2018).

Classically, the *in vitro*-derived M1/M2 nomenclature has been used to describe macrophage phenotypes upon SCI (Nahrendorf & Swirski, 2016). Briefly, acute phase macrophages have been described as adopting the M1, pro-inflammatory, cytotoxic, phenotype, while the subacute and chronic phase macrophages adopt M2, anti-inflammatory, regeneration-promoting states (Kong & Gao, 2017; Milich et al., 2019). Albeit the opposite trend has also been argued (X. Wang et al., 2015). Whether these simplified labels reflect the true states of macrophages upon SCI, in both animal models and humans, has yet to be elucidated. However, a recent single-cell RNA sequencing (scRNAseq) study of CNS cells upon SCI was unable to find evidence for this classification (Milich et al., 2021), in line with studies from other CNS pathologies (Kim et al., 2016; Ransohoff, 2016). The historic challenge of differentiating activated macrophages and microglia adds to the uncertainty of whether the M1/M2 classification remains relevant (Alizadeh et al., 2019).

Polarisation aside, several macrophage functions after SCI are well established. They have been demonstrated to further activate astrocytes and microglia and contribute to the inflammation by secreting cytokines and complement components that maintain the permeabilization of the BSCB (Alizadeh et al., 2019; Kong & Gao, 2017). Based on findings in other disease models and *in vitro* studies, this pro-inflammatory state is likely fuelled by a

metabolic shift from oxidative phosphorylation (OXPHOS), seen in homeostatic and anti-inflammatory macrophages, to aerobic glycolysis. This shift is known in cancer biology as the Warburg effect (Bernier et al., 2020; Devanney et al., 2020; Peruzzotti-Jametti et al., 2021; Warburg, 1956). Macrophages also represent a source of trophic and anti-inflammatory molecules that promote regeneration (Brennan & Popovich, 2018; Popovich & Jones, 2003). Macrophages have been shown to injure oligodendrocytes in the lesion epicentre through phagocytosis and cytotoxic cytokine production leading to their apoptosis, and modulate the adaptive immune response, thereby facilitating the immunological demyelination by myelin-derived antigen presentation through MHCII (Alizadeh et al., 2019; David & Kroner, 2011; Kong & Gao, 2017). Thus, macrophages represent dynamic players in the SCI pathology, but we currently lack a broader understanding of how each of these functions are related to each other in time and space across the SCI pathology.

CNS-Associated Macrophages

The recently characterised CAMs, also known as border-associated macrophages, are long-lived non-parenchymal CNS resident macrophages, most of which share the ontogeny of microglia (Goldmann et al., 2016; Jordão et al., 2019; Ransohoff & Cardona, 2010; Van Hove et al., 2019). Very recently, these cells were observed to infiltrate the CNS parenchyma in a bone marrow chimera fate-mapping model of SCI (Bellver-Landete et al., 2019). The authors found that CAMs infiltrated the injured spinal cord and contributed minimally to the microglia scar at 14 dpi. Whether chimeras were suitable for this type of study is unclear, as irradiation compromises blood-brain barrier (BBB) and BSCB integrity and alters the function of resident myeloid cells (Bruttger et al., 2015; Mildner et al., 2011; Mildner et al., 2007). A recent scRNAseq study of all CNS infiltrating and resident cells after SCI identified small quantities of CAMs at 1, 3, and 7 dpi (Milich et al., 2021). However, all spatial context is lost through single cell isolation and this study isolated the spinal cord through dissection, *vs* hydraulic extrusion, making it unclear whether the observed CAM population had infiltrated the parenchyma. CAMs have been demonstrated to play vital roles in other CNS pathologies, thus it's likely that these cells do play an active role in the SCI pathology that has yet to be elucidated (Kierdorf et al., 2019).

MICROGLIA

Microglia are the resident immune cells of the CNS and share many phenotypes and functions with macrophages[†] throughout the SCI pathology (Alizadeh et al., 2019; Li & Barres, 2018) and as with macrophages, the relevance of the M1/M2 nomenclature to describe microglia activation is a matter of debate, with many articles trending away from the use of this over-simplification (Ransohoff, 2016). Like macrophages, microglia have been described as both pro- and anti-inflammatory after SCI. Studies in other disease contexts have demonstrated that human pro-inflammatory microglia become glycolytic, as seen in cancer cells, and more recently, pro-inflammatory macrophages (Hasselmann et al., 2019; Peruzzotti-Jametti et al., 2021; Warburg, 1956). While microglia metabolism after SCI is not well understood, the evidence for this metabolic switch suggests an interesting role for microglia metabolism in modulating the immune response and represents an under investigated therapeutic avenue.

In contrast to macrophages, microglia respond earlier to the pathology, given their proximity to the lesion site (Alizadeh et al., 2019; Bowes & Yip, 2014). Microglia, but not macrophages, have also been described to migrate to the borders of the lesion and proliferate extensively during the first week post-injury in mouse models of SCI (Bellver-Landete et al., 2019; Greenhalgh & David, 2014; Milich et al., 2021; Noristani et al., 2017). Here, they also form a dense cellular border, known as the microglia scar, between infiltrating cells and scar-forming astrocytes, which works to compact the lesion epicentre (Bellver-Landete et al., 2019; Zhou et al., 2020). The PLX5622-induced depletion of microglia and CAMs during the first week post-SCI disrupted glial scar formation and worsened functional outcomes (Bellver-Landete et al., 2019). Whether this corralling function occurs in the human condition has yet to be investigated.

Like macrophages, microglia play a role in phagocytosing cellular debris, including myelin debris (Brennan & Popovich, 2018). However, the degree to which microglia *vs* macrophages scavenge myelin debris is under debate. Microglia have been reported to phagocytose myelin debris through a TREM2-mediated pathway (Kopper & Gensel, 2018) and have been proposed

[†] While microglia are a type of macrophage, this dissertation will reserve the term '*macrophage'* for haematopoietic CNS infiltrating macrophages

as the primary scavengers of myelin debris in the acute phase of SCI (Greenhalgh & David, 2014). This role is thought to be taken over by infiltrating macrophages from 3 dpi. Furthermore, there are reports that macrophages perform better at scavenging myelin debris, compared to microglia, hypothesised to be due their expression of lipid-sensing receptors, MAC2 (*Lgals3*), ABCA1, and MSR1 in macrophages (Kopper & Gensel, 2018; X. Wang et al., 2015). Yet, activated microglia have also been demonstrated to express these genes in mouse models of SCI (Milich et al., 2021). Unlike macrophages, there are no reports of microglia adopting foam-cell phenotypes and persisting to the chronic phase of SCI. Whether this reflects the biology, or a historic lack of appropriate tools to differentiate macrophages from microglia, has yet to be determined.

In summary, the immune response to SCI is highly dynamic, involving many heterogenous CNS resident and infiltrating cell types with temporally specific roles that are highly dependent on environmental cues. The immune response comprises many beneficial functions, with just as many detrimental ones, thus unsurprisingly, therapies that aim to impede inflammation have been largely unsuccessful (Schwab et al., 2014). A better understanding of the cell-specific functions, with temporal and spatial acuity, is needed to consolidate our fragmented understanding of the SCI pathology.

1.2 Animal Models

Given the complex pathophysiological response to SCI, it is not currently possible to accurately model the pathology in its full complexity *in vitro* or *in silico*. Studying SCI in human tissue is constrained by the scarcity of samples, the inability to perform controlled, hypothesis-driven experiments, and the commonly used formalin-fixed paraffin-embedded (FFPE) preservation method, which, at least historically, has presented challenges for molecular-level research. As such, researchers have turned to animal models of SCI to better understand the pathology, with the ultimate goal of identifying putative therapeutic approaches for this debilitating condition.

1.2.1 Injuries

After decades of research, there are now many well-established, reproducible animal models of SCI. The most common models include contusion, compression, transection, distraction,

and dislocation. Each of these models work to either recapitulate the human condition or to isolate a specific element of the pathology for further investigation (Cheriyan et al., 2014; N. Zhang et al., 2014).

CONTUSION

Contusion SCI is caused by blunt force impact to the spinal cord, typically after laminectomy but without the penetration of the meninges (Marinelli et al., 2019; Tyler et al., 2013). The severity of the lesion (mild, moderate, severe) can be easily and reproducibly controlled based on the applied force. This injury directly affects the grey matter and recapitulates the biomechanics and many of the secondary injury dynamics that occur in the most prevalent form of the human pathology (Sharif-Alhoseini et al., 2017). Thus, for studies looking to investigate the disease pathology in its entirety, the contusion model currently represents the most appropriate option. To ensure reproducibility, contusion injuries are generated using clamps to secure the spinal column and computerised impactors that record various parameters depending on the device (Cheriyan et al., 2014).

Very recently, a novel contusion model was proposed in which a laminectomy is not performed, and a PinPoint precision impactor device (Stoelting) exerts pressure directly on the bone (Marinelli et al., 2019). This approach claims to model the human contusion SCI condition more accurately, and notably addresses two major caveats of the former approach: the formation of cystic cavities and the complete and persistent absence of motor recovery until at least 30 dpi in mouse models. This model is promising but requires further investigation and independent replication before replacing the classic contusion model.

Most contusion models induce thoracic-level injuries, to minimise animal suffering, however, cervical-level injuries are more clinically relevant, as cervical injuries in humans represent almost 50% of cases (Alizadeh et al., 2019; Cheriyan et al., 2014; Sharif-Alhoseini et al., 2017). The different injury levels result in distinct functional deficits and subtly different pathological responses based on the varying degree of white *vs* grey matter damage (Dunham et al., 2010). For the investigation of the endogenous pathology dynamics after SCI, the trade-off between animal welfare and clinical relevance is perhaps on the side of the thoracic injury. However,

for pre-clinical investigations of novel therapeutics, cervical contusion injuries would better reflect the clinical relevance of the study outcomes (Cheriyan et al., 2014).

Compression

Compression SCI models are generated by applying persistent pressure to the spinal cord. This can be achieved through several different mechanisms, the most clinically relevant of which are modified aneurysm clips because they also comprise an element of contusion (Cheriyan et al., 2014). Other approaches include calibrated compression forceps, balloon compression, or more recently, spinal cord strapping (Cheriyan et al., 2014). Each of these approaches bid for reproducibility by applying the compression at a measurable value (e.g., force or displacement), for a set period of time and are relatively accessible and easy to generate as they do not require a computerised device like many contusion methods. Like contusion, compression injuries aim to recapitulate the entire pathology. The clinical relevance of this model stems from its parallels with traumatic injuries, namely fracture dislocations and burst fractures (Cheriyan et al., 2014; Sharif-Alhoseini et al., 2017).

Transection

After a laminectomy, the transection model is generated by making a surgical incision to sever the spinal cord either bilaterally (full transection) or unilaterally (partial transection) (N. Zhang et al., 2014). This model has minimal clinical relevance as transections and lacerations of the spinal cord of humans are rare. However, it is a valuable model for studying axonal regeneration, functional outcomes, and tissue engineering approaches. These models are also easy to generate, and in the partial transection model, the same animal can act as its own control. Furthermore, this model results in better animal welfare and a reduced need for post-operative care compared to the contusion and compression models (Cheriyan et al., 2014). Unfortunately, given that the model is manually generated, it can have mixed reproducibility, especially in the partial transection model.

Distraction

The goal of distraction injury models is to simulate the biomechanics of the tension force inflicted upon the cord in distraction traumatic SCI but also inadvertent injuries that occur during spinal surgeries (Seifert et al., 2011). To generate this model, a computerised device is used to apply opposing traction forces, stretching the spinal cord (Sharif-Alhoseini et al.,

2017). The severity of the injury is a function of the distraction length. This model is not as well characterised as the former three models, but prospectively represents a promising tool (Cheriyan et al., 2014).

DISLOCATION

Investigations into the pathological response to dislocation injuries, which are not entirely uncommon forms of traumatic SCI, are vastly understudied. Equally, the dislocation model is underdeveloped in comparison to the other models (Cheriyan et al., 2014). Biomechanically, the model aims to laterally displace select vertebra, inflicting injury primarily in the white matter (Choo et al., 2009; Sharif-Alhoseini et al., 2017). However, this model requires further characterisation as well as the establishment of standards for reproducibility.

1.2.2 SPECIES

The SCI anatomy and pathology is largely, but not perfectly, conserved across mammals (Courtine et al., 2007; N. Zhang et al., 2014). Understanding the species-specific or even strain-specific differences is key to selecting an appropriate model for the scientific aim or question (Fouad et al., 2013) (Figure 1.4). Rodents represent the most commonly used species (90% of studies) followed by cats (2.3%), dogs (2.2%), non-human primates (1.5%) and pigs (1.5%), and finally non-mammalian vertebrates (1%) (Sharif-Alhoseini et al., 2017).

MICE

Mice models of SCI account for approximately 16% of animal SCI models (Sharif-Alhoseini et al., 2017). These models are easy and affordable to handle and house, and their high reproductive rates leave them in no short supply and make them good candidates for transgenic studies (Figure 1.4). Mice have similar genomes to humans and have been well-characterised both on a genetic and anatomical level (Courtine et al., 2007). For this reason, as well as the aforementioned advantages, mice are becoming increasingly popular for cellular and molecular-level studies (Sharif-Alhoseini et al., 2017). However, in the context of SCI, mice, and rodents in general, differ from humans in several ways.

First, mice and humans differ in their spinal anatomy. The most obvious difference is the corticospinal tract (CST). The tract itself remains dorsal in rodents, whereas in primates it has migrated ventrally (Courtine et al., 2007). The projections of the rodent CST are much less

complex than primates, with projections limited to the dorsal horn and premotor spinal circuits (Courtine et al., 2007). This anatomical difference is reflected in the decreased level of dexterity observed in rodents. Also, in rodents, damage to the CST alone does not disrupt locomotion, unlike in primates. Similarly, the anatomy of autonomic innervation in rodents is slightly different than primates with bladder control and sexual functions being the main discrepancies. Another important difference is the diameter and length of the spinal cord. Evidently, the mouse cord is much shorter and narrower than the primate cord, which has important implications for the regeneration and successful reinnervation of axons (Akhtar et al., 2008; Courtine et al., 2007). Ultimately, these anatomical differences are particularly pertinent for pre-clinical studies assessing regeneration and functional outcomes.

Second, despite being the primary model for immunological studies in general, the leukocyte populations, and subsequently the innate and adaptive immune responses, are not identical between mice and humans (Mestas & Hughes, 2004). For example, mice have a greater proportion of lymphocytes and lower proportion of neutrophils in their blood. Mice macrophage gene products have several differences that likely impact the innate immune response, such as Toll-like receptor and FcR identities and increased nitric oxide production in response to inflammatory stimuli (Mestas & Hughes, 2004). Mice and humans express different immunoglobulin isotopes and cytokines conserved between the two species can induce different immunoglobulin class switches. This is by no means an exhaustive list of differences. However, these exemplify the inter-species discrepancies, the full effects of which are unknown in SCI, and may help to justify the temporal differences in peripheral immune cell infiltration observed between mice and humans (Bowes & Yip, 2014). Most strikingly, mouse lymphocytes, including B-Cells, infiltrate the spinal cord in a biphasic manner, the first wave beginning around 7 dpi and peaking around 14 dpi and the second wave peaking around 6 weeks post-SCI (Bowes & Yip, 2014; Donnelly & Popovich, 2008). In humans, T-cells cells begin around 2-3 dpi and peak around a month post-SCI, with no biphasic trend observed. Additionally, based on immunohistochemistry (IHC) staining of B-Cell marker CD20, B-Cells have not been observed in the human injured spinal cord parenchyma (Bowes & Yip, 2014; Donnelly & Popovich, 2008; Fleming et al., 2006). Conversely, myeloid cell temporal dynamics appear to be well-aligned across rodents and humans (Fleming et al., 2006; Sroga et al., 2003).

To account for the inter-species leukocyte discrepancies, recent studies have turned to humanised mouse models to investigate the immune response to SCI (Carpenter et al., 2019; Carpenter et al., 2015). In these experiments, immunocompromised mice are reconstituted with human immune cells before inflicting SCI. The pathological and functional outcomes of these humanised mice are worse than their non-engrafted counterparts (Carpenter et al., 2019). However, the reproducibility of this model and whether the observed differences are clinically relevant remains to be seen.

A final notable difference between the human and mouse SCI pathology is formation of the cystic cavity, which typically does not occur in the mouse model. The lack of the cyst is hypothesised to be due to an enhanced T-cell response and the increased fibrosis in the mouse lesion (Fleming et al., 2006). There are several exceptions, including the CD1 mouse strain (Inman et al., 2002) or the novel bone-impacting contusion model (Marinelli et al., 2019), demonstrating the modest variability of the pathology between strains and injuries. Ultimately, mouse models of SCI remain a useful tool, but fail to fully recapitulate the human condition and are thus best suited for basic and early pre-clinical studies.

RATS

Rats are the most common animal model of SCI, representing 72% of all studies (Sharif-Alhoseini et al., 2017). Like mice, they are cost-effective and easy to use and house (Figure 1.4). They are well-characterised on an anatomical level, years of SCI research has resulted in well-established injury protocols, and numerous functional recovery paradigms have been described to evaluate rat models of SCI (Courtine et al., 2007; J. J. Li et al., 2020; Sharif-Alhoseini et al., 2017). Unlike mice, they are not as readily available for transgenic studies (Ellenbroek & Youn, 2016). Similar to mice, the CST and general size of the rat spinal cord differs from humans. The infiltration of lymphocytes in rat SCI differs from both the human and mouse pathology, with rat lymphocytes infiltrating around 3-7 dpi alongside monocytes (Donnelly & Popovich, 2008). Furthermore, unlike mice, rats develop the cystic cavities observed in the human condition (Li et al., 2020). Ultimately, rats are appropriate models for preliminary studies of the SCI neuropathology, but their caveats (namely barriers to transgenics and their anatomical and pathophysiological deviations from human SCI) should

be taken into consideration and, if feasible, complimented with non-invasive human studies to ultimately improve the translability of these findings into the clinic.

NON-HUMAN PRIMATES

The use of non-human primates for SCI research accounts for only 1.5% of studies (Sharif-Alhoseini et al., 2017). This is largely due to the high costs and immense ethical concerns that surround these animal models (J. J. Li et al., 2020) (Figure 1.4). However, even the small new world primates like squirrel monkeys and marmosets more accurately model the human condition compared to rodents (Cheriyan et al., 2014). This can be attributed to a greater similarity to the human genome, anatomy, and spinal cord circuitry (J. J. Li et al., 2020). Larger, old world non-human primates also recapitulate the autonomic aspects of the pathology, including bladder and sexual functions, as well as the regulation of blood pressure (Courtine et al., 2007). These larger models are also important for the final pre-clinical testing of tissue engineering and surgical approaches before their use in clinical trials (J. J. Li et al., 2020). Thus, while non-human primate models of SCI are valuable, they come at a great cost to animal welfare and should be strictly reserved for late pre-clinical studies that cannot be accomplished in lower species.

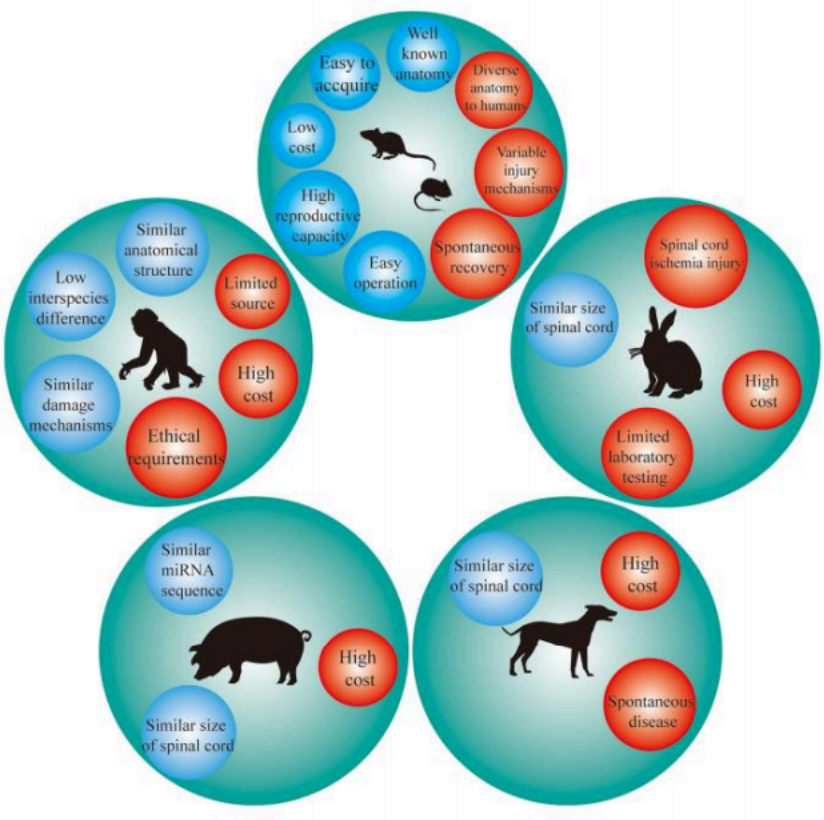


Figure 1.4 The pros (blue) and cons (red) of the most common animal models of SCI. No animal model is perfect and thus careful consideration must be made to select the optimal model for the scientific goal or question. Figure from (J. J. Li et al., 2020)

NON-MAMMALIAN VERTEBRATES

Non-mammalian vertebrates, such as fish and amphibians, account for about 1% of all SCI studies (Sharif-Alhoseini et al., 2017). While these models do very little to simulate the human condition, they are useful for understanding regeneration on a very basic level as they are able to fully recover from SCI (Lee-Liu et al., 2013). Critically, these regenerative model organisms have provided better insight into the regenerative nonpermissive microenvironment of the mammalian system, with features such as glial, inflammatory, and immune responses highlighted as key gatekeepers of regeneration (Lee-Liu et al., 2013).

1.3 SINGLE-CELL TRANSCRIPTOMICS

Transcriptomics can be defined as the investigation of the RNA transcripts produced by the cell from a genomic DNA template via high-throughput methods. This approach is typically employed as a substitution for studying the proteome, high-throughput full-scale methods for which are on the horizon (e.g. via mass spectrometry), but are not yet readily available (Haque et al., 2017). The advent of transcriptomics, including microarrays and bulk RNAseq, revolutionized our understanding of gene expression and led to unprecedented breakthroughs in biology and medicine (Picelli, 2017). However, such approaches fail to characterize the transcriptional heterogeneity of individual cells, a major caveat given that cells of the same cell type[‡] can occupy different cell states, measured as biologically relevant differences in both transcript identity and quantity (Papalexi & Satija, 2018). Single-cell transcriptomics, achieved through scRNAseq, can elucidate cellular heterogeneity at an unprecedented resolution, even in seemingly homogenous populations (Figure 1.5).

[‡] The task of defining *cell type* vs *cell state*, particularly in the single cell era, is not an easy one (Morris, 2019; Trapnell, 2015). However, for the purposes of this dissertation, I have reserved the term *cell type* for canonical categories, such as microglia or neutrophils. Meanwhile, I have defined *cell states* as the condition of a cell compared to others within its cell type, for example homeostatic *vs* pro-inflammatory microglia. The arbitrary nature of these definitions is acknowledged.

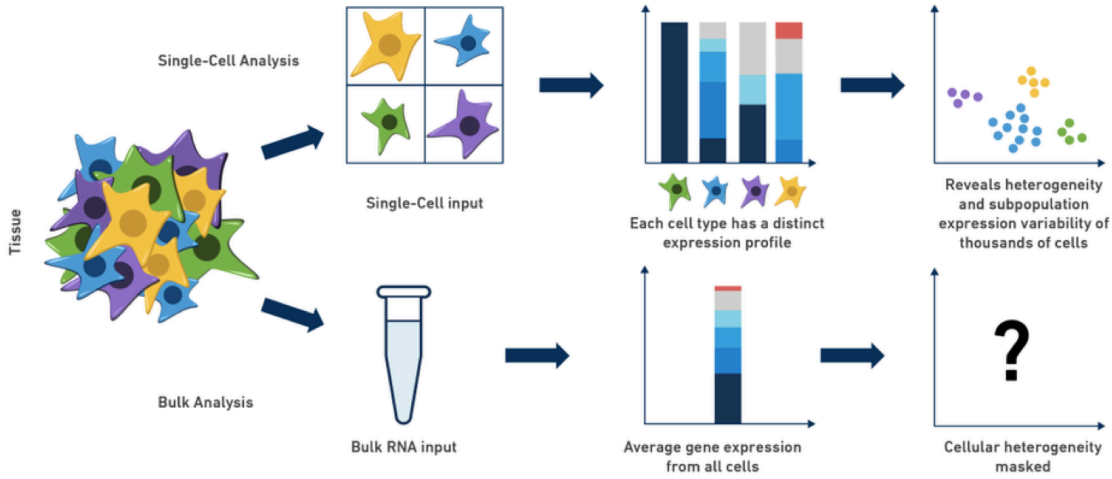


Figure 1.5 A diagram demonstrating the power of scRNAseq to elucidate heterogeneity in biological samples. Compared to bulk RNAseq, scRNAseq unveils cellular level heterogeneity. Figure From: (Sheila-10X, 2017)

1.3.1 Generating High-Throughput Single-Cell Transcriptomics Data

To generate high-throughput single-cell transcriptomics data, a single-cell suspension of viable, high-quality cells must be prepared. For liquid tissues, such as blood, relatively rapid and straightforward techniques like the generation of a density gradient can be applied (Reichard & Asosingh, 2019). Conversely, for *in vitro* samples and even more so, *ex vivo* solid tissue samples, cells must be dissociated. This can be achieved through mechanical or enzymatic methods. The most appropriate approach depends on the sensitivity and fragility of the cell type, the composition of the extracellular matrix, and other experimental design considerations, such as the integrity of cell surface-markers (Reichard & Asosingh, 2019). After dissociation, the suspension is filtered to remove poorly dissociated tissue or cell aggregates and, in the case of CNS tissue, myelin debris must also be removed. Finally, cell viability must be assessed and dead or dying cells must be removed. Typical approaches include commercially available “clean-up” kits or the addition of a live/dead cell dye before fluorescent activated cell sorting (FACS).

There are two main approaches to isolate individual cells from a single-cell suspension, droplet-based and plate-based (Figure 1.6). Both of these methods capture and barcode poly-adenylated RNA (namely mRNA) for downstream identification via next generation sequencing (NGS). Droplet-based approaches, such as inDrop (Klein et al., 2015) or 10X Chromium (Zheng et al., 2017), rely on microfluidics to entrap individual cells from the single-cell suspension into nanolitre size droplets that contain the reagents for the generation of cell

libraries through reverse transcription and transcript barcoding (Griffiths, Scialdone, et al., 2018; Haque et al., 2017; Ziegenhain et al., 2017). Many droplet-based approaches, such as 10X Chromium, barcode transcripts with short (~16 base pair [bp]) cell barcodes, the sequence of which is unique to each droplet, and shorter (~10bp), randomly generated nucleotide sequences, called unique molecular identifiers (UMIs). The latter allows amplified transcripts to be deduplicated, i.e., traced back to their original mRNA, enabling the user to discriminate between transcripts that were highly expressed *in vivo vs* those that were highly amplified *ex vivo.* In this way, UMIs also allow mRNA transcripts to be directly quantified. After each transcript is barcoded, the cell libraries are then pooled and prepared for NGS in the fashion of bulk RNAseq.

The droplet-based methods are relatively low cost and incredibly high-throughput ($1x10^3$ – $1x10^6$ cells) making them ideal for sampling diverse cell populations (Griffiths, Scialdone, et al., 2018; Ziegenhain et al., 2017). The trade-off is smaller cell libraries; this approach is only able to determine the relative volume of moderately/highly expressed genes, which might be a concern if certain weakly expressed genes are of key interest to the researcher (Haque et al., 2017). Notably, this sensitivity is increasing as these protocols are refined (Yamawaki et al., 2021). Most of these approaches also use a 3’ or 5’-end counting protocol, which does not provide full-length transcript data, making certain inquiries like splice-variant analysis, infeasible (Haque et al., 2017), however solutions to this challenge are approaching (Mamanova et al., 2021). Finally, droplet-based methods are more likely to generate multiplets (i.e., two or more cells captured and barcoded within a single droplet). However, many computational approaches have been developed to mitigate this during the analysis (Amezquita et al., 2020). Samples can also be labelled with oligonucleotide-barcoded lipids (McGinnis et al., 2019) or antibodies (Stoeckius et al., 2018) allowing them to be multiplexed during the droplet-based isolation step. This technique reduces batch effects between samples, allows for multiplet identification, and enables users to load and therefore capture more cells per run (Stoeckius et al., 2018). Ultimately, despite its caveats, droplet-based protocols are extremely popular and effective for evaluating heterogeneity in complex systems.

Plate-based methods, such as MARS-seq (Jaitin et al., 2014) or Smart-seq3 (Hagemann-Jensen et al., 2020), require cells to be sorted into microwell plates, often via fluorescent activated cell sorting (FACS), where they are then lysed, barcoded, and amplified in the well before NGS (Griffiths, Scialdone, et al., 2018; Haque et al., 2017). Notably, some plate-based methods (e.g., MARS-seq, CEL-seq2) allow for the cost-effective pooled PCR amplification and library preparation utilised in droplet-based methods (Papalexi & Satija, 2018). Plate-based methods are modestly high-throughput ($1x10^2$ – $1x10^3$) and, in labs where robotic automation is unavailable, tend to be lengthy and manually intensive (Papalexi & Satija, 2018). However, these approaches do typically result in extremely high sensitivity, allowing for the detection of lowly-expressed genes (Haque et al., 2017). Thus, they are typically employed to profile limited samples, like rare cell types during development (Mamanova et al., 2021).

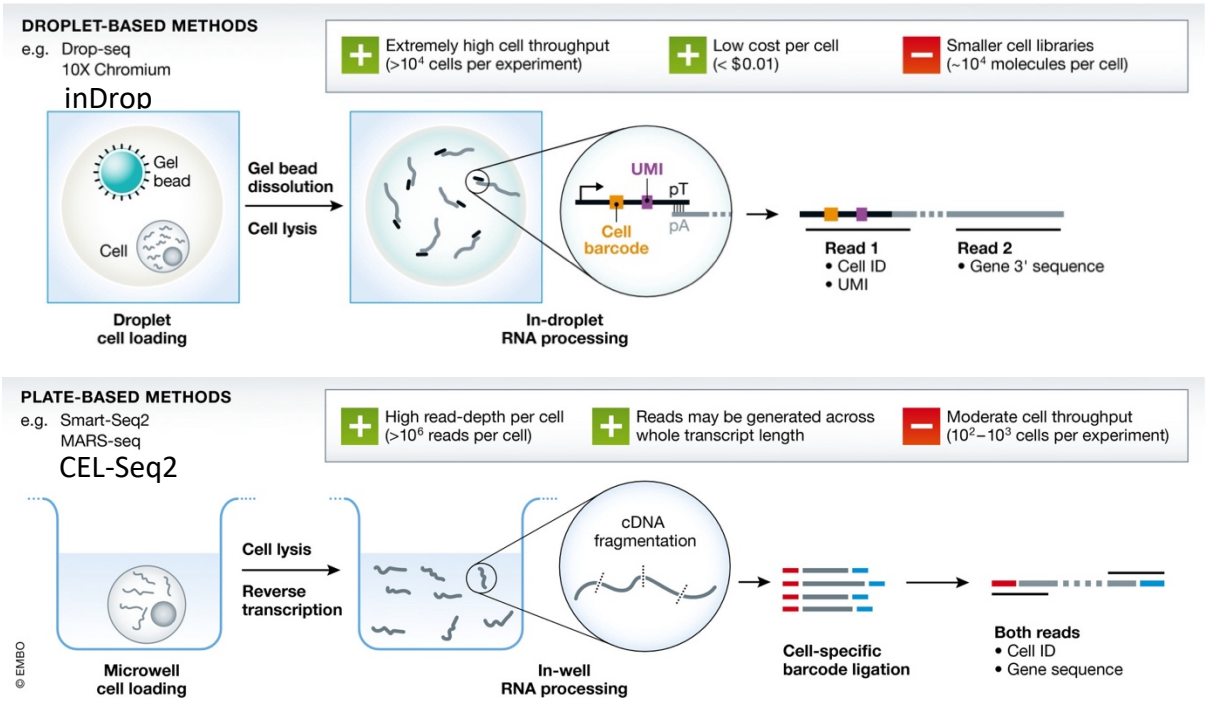


Figure 1.6 The two main approaches to single-cell isolation for generating single-cell transcriptomics data. Droplet-based approaches are generally higher throughput at the cost of library size. Note that the schematic does not include each step in the protocol, but rather highlights those that differ between the two methods. Figure adapted from (Griffiths, Scialdone, et al., 2018).

1.3.2 A Brief Introduction to the Handling of scRNAseq Data

scRNAseq generates immense quantities of data that present unique computational challenges. The methods and considerations required for handling scRNAseq data will be discussed in great depth throughout this dissertation. However, to get started I will briefly outline the key steps.

NGS typically produces a data file in the binary base call format, which records the identity, location, and confidence level for each detected fluorophore on the flow-cell during NGS. This file must be converted to a demultiplexed FASTQ format and aligned to the genome of the applicable species. Then low-level analysis can be performed. This typically includes cell-calling (i.e., distinguishing cell-containing *vs* empty droplets), quality control (QC) including multiplet removal, normalisation, highly variable gene (HVG) selection, dataset integration and batch correction if required. The result is a cleaned feature-barcode matrix that is ready for higher-level analysis and interpretation.

Approaches to facilitate biological interpretation and discovery through scRNAseq are flexible and numerous. However, a fundamental approach is dimensionality reduction. In the context of scRNAseq data, dimensionality reduction is the process of representing the expression of thousands of genes in just a few dimensions (Amezquita et al., 2020). A typical approach for dimensionality reduction is principal component analysis (PCA). PCA is a deterministic technique that considers the high-dimensional space and determines the linear projection, or principal component (PC), that captures the greatest amount of variance (Amezquita et al., 2020; Rudi et al., 2007). The subsequent PC must be orthogonal to the previous and describe the greatest variance out of that which remains unaccounted for. The resultant PCs are ordered by decreasing proportions of the variance explained and typically the top 10-50 PCs are selected for downstream analysis (Amezquita et al., 2020). Retaining only the top PCs allows for more efficient downstream processing and helps to reduce technical noise, as the biological heterogeneity is expected to drive most of the variance in the dataset.

The high-dimensional gene expression of each cell can be further condensed into 2-3 dimensions for visualisation purposes. Plotting the top two PCs would accomplish the task. However, as mentioned, PCA calculates linear distances, making it poorly suited for visualising complex populations (Figure 1.7A, B). Alternative approaches, such as diffusion maps, t-stochastic neighbour embedding (t-SNE), and uniform manifold approximation and projection (UMAP) are not constrained by linear relationships. The diffusion map method has been adapted for scRNAseq data (Haghverdi et al., 2015) and can be interpreted as representing the data based on predicted transitions between cells (Figure 1.7C). Thus, it is well suited for

visualising differentiating cells, such as those observed during development. UMAP and t-SNE are two fundamentally different algorithms that share several features, including their stochastic and non-linear nature and their goal of conserving the relationships between cells that are similar in the high-dimensional gene expression space (Amezquita et al., 2020). Unlike PCA and diffusion maps, UMAP and t-SNE do not preserve the relationships beyond the nearest neighbours in high dimensional space. This means we can only draw conclusions about the local, but not global structure of the map (Figure 1.7D) (Amezquita et al., 2020; van der Maaten, 2008).

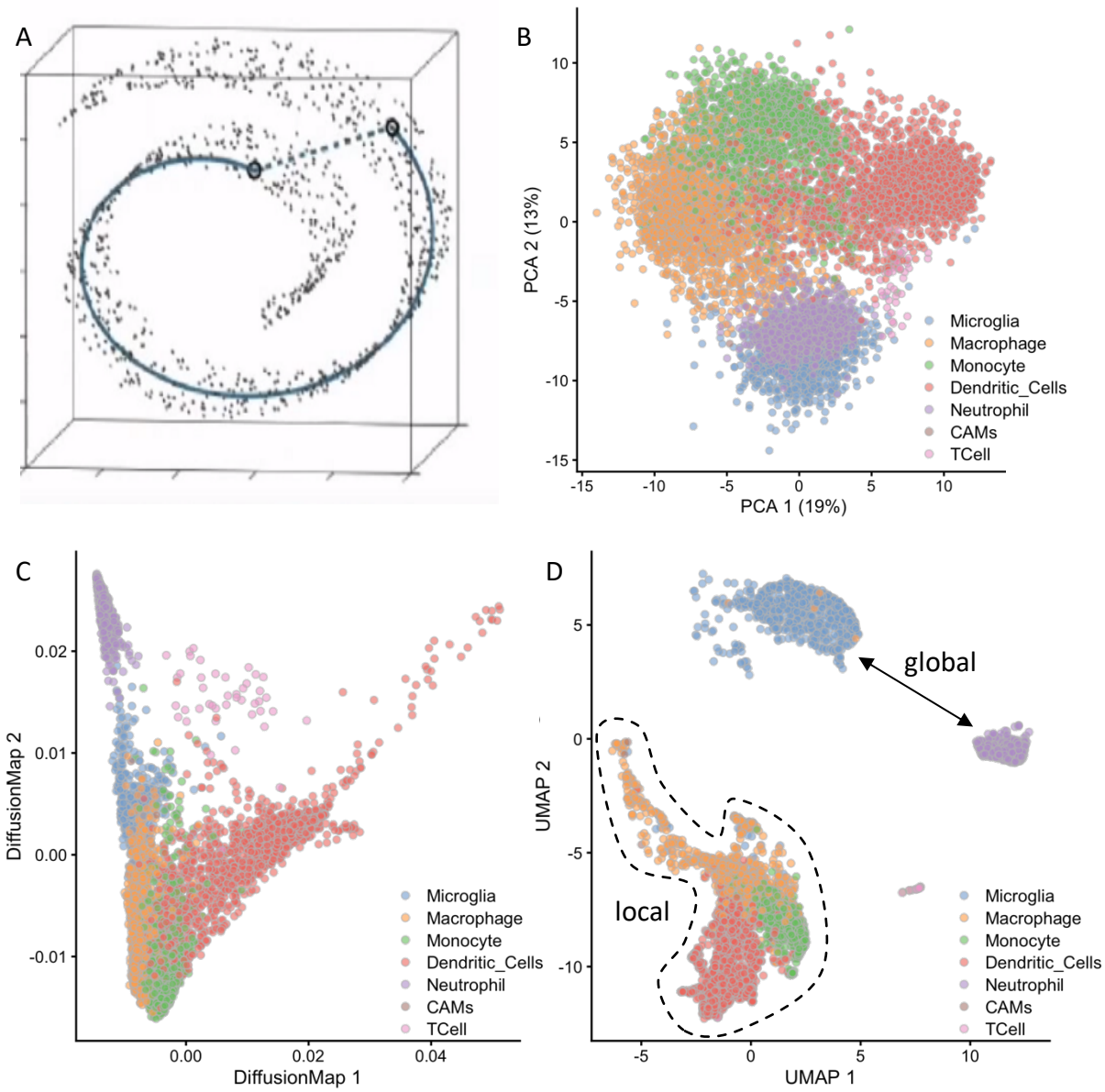


Figure 1.7 A) The relationships between cells as generated by PCA (dotted line) vs non-linear dimensionality reduction approaches (solid line). Image from (van der Maaten, 2013), discussing PCA vs t-SNE. B) A subset of this myeloid cell atlas visualised using PCA. Each cell is coloured by its annotated cell type. C) The same data as a diffusion map. C) The same as a UMAP. Local vs global structure is demonstrated.

1.3.3 High-Throughput Transcriptomics in SCI

High-throughput transcriptomic analyses of the SCI pathology have undoubtably broadened our understanding of the pathology in animal models and have generated many new hypotheses, inspiring or complimenting mechanistic studies. Many of these studies have been generated through bulk RNAseq, with a few scRNAseq approaches published within the last decade. Notably, no spatial transcriptomics studies have been performed to date, despite the spatial nature of the SCI pathology. In this section, I will briefly discuss some of the high-throughput transcriptomics SCI studies, highlighting their contributions and their caveats, and linking their findings to complimentary, non-transcriptomics studies.

Bulk RNAseq Studies

Several studies have employed bulk RNAseq to investigate transcriptional changes of neural tissue upon SCI. Here I will discuss a few key publications that have deepened our understanding of the SCI pathology, or have taught us valuable lessons about the optimal design of transcriptomics studies in SCI.

A study by Chen et al., 2013 used thoracic contusion mouse models of SCI to sequence the lesion epicentre (0.5 mm of tissue) in sham injured controls and at 2 and 7 dpi. They calculated differentially expressed genes (DEGs) between conditions and then performed gene set enrichment analysis. The top enriched pathways included the expected inflammatory response, cell death, CNS development, but also atherosclerosis signalling, supporting a role for foamy macrophages after SCI (Fleming et al., 2006; X. Wang et al., 2015; Zhu et al., 2017). By sequencing RNA from all cell types they were able to investigate global pathways but their technique did not provide an understanding of individual cell types, whether a pathway is driven by a single cell type, the nature of the cellular interactions driving the observed pathway, or the spatial context of these proposed functions (Chen et al., 2013). This study was repeated in rats (Shi et al., 2017) and marmosets (Nishimura et al., 2014), providing global gene expression resources, but faced the same pitfalls as Chen et al., 2013.

In another study, GFAP-RiboTag mice (Sanz et al., 2009) were used to isolate and sequence the astrocyte translatome at 14 dpi in mouse thoracic compression models of SCI (Anderson

et al., 2016). Unlike the aforementioned global sequencing studies, this project allowed for an in-depth investigation into a specific cell type and was paired with a functional study, providing both a descriptive and mechanistic understanding of astrocytes and the astrocyte scar after SCI. Specifically, they found that interrupting astrocyte scar formation failed to promote regeneration, in line with previous findings in rat models of contusion SCI (Okada et al., 2006). The RiboTag results demonstrated that this could be due to the fact that astrocytes within the lesion epicentre secreted axon-growth supporting molecules (Anderson et al., 2016). RiboTag was a particularly appropriate approach to study the gene expression of astrocytes after SCI as these cells are extremely fragile and are not easily dissociated from homeostatic tissue, let alone thick, fibrous scar tissue. However, the authors could not draw conclusions on the heterogeneity of astrocytes after SCI (Anderson et al., 2016).

Similarly, LysM(*Lyz2)*-RiboTag was used to isolate *Lyz2*$^{+}$ cells, namely macrophages and a subset of microglia, at 3 and 7 dpi in mouse models of thoracic contusion SCI (Zhu et al., 2017). They performed Gene Ontogeny (GO) enrichment analysis and KEGG pathway analysis and found that acute phase *Lyz2*$^{+}$ cells are characterised by migration and cytokine production, while by 7 dpi the primary role of these cells is to catabolize lipids, again supporting a role for foamy macrophages after SCI (Fleming et al., 2006; X. Wang et al., 2015; Zhu et al., 2017). The RiboTag approach also meant that this phenotype was not driven by RNA in phagocytic cargo or isolation-induced transcription (Haimon et al., 2018). The authors performed a complementary experiment in which they used IHC to investigate post-SCI macrophages in a CD36 global knock out (KO) mouse (Zhu et al., 2017). They found that by 7 dpi, macrophages in the wild-type (WT) mice were lipid-laden, while in the KO this was substantially decreased. Furthermore, they observed smaller lesion sizes and modestly improved functional outcomes in the KO mice compared to WT mice. Cumulatively this study replicated the finding of foamy macrophages after SCI, but the bulk sequencing and global KO meant the authors were unable characterize the heterogeneity of these cells and determine whether all or merely a subset of *Lyz2*$^{+}$ cells adopted this foam cell-like phenotype. Furthermore, it's unclear if *Lyz2*$^{-}$ cells also take on this state.

Cx3cr1^{+} microglia were recently sequenced from 10 mm of tissue in mice models of thoracic transection models of SCI at 3, 7, and 14 dpi and uninjured controls (Noristani et al., 2017).

Cells were isolated via enzymatic dissociation followed by FACS and it appears that most of the protocol was performed at room temperature or 37°C. The authors used a panel of non-microglia myeloid cells markers to attempt to enrich the *Cx3cr1*$^{+}$ population for microglia. However, most of these markers have since been demonstrated to become upregulated in activated microglia (Amici et al., 2017). This suggests that this study enriched their samples for homeostatic-like microglia, which would be present away from the site of transection, but well within the large tissue section they collected, emphasising the importance of spatial context in SCI. Unsurprisingly, the authors found that only the 3 dpi sample differed drastically from the control. They also argued that microglia upregulate the astrocyte marker, *Vim*. The expression of *Vim* has been reported in activated microglia (Jiang et al., 2012), but in the Noristani et al., 2017 dataset it was likely ambient mRNA. Finally, Noristani et al., 2017 claimed that microglia upregulated *Brca1* after SCI, signifying DNA damage. However, this finding was not confirmed via tissue pathology in mice and has not been replicated; it likely represents a technical artefact induced by their isolation protocol. This study highlights the importance of careful experimental design and validation of key findings through complementary methods, such as tissue pathology.

A very recent study sequenced myeloid samples from laminectomy-only controls and 3, 7, and 14 dpi. To isolate these cells for bulk RNAseq, they used isolation of nuclei tagged in specific cell types (INTACT) mice crossed with *Cx3cr1*^{CreER} mice to isolate myeloid lineage nuclei from 10 mm of tissue in mouse models of thoracic transection SCI (Zhou et al., 2020). The INTACT method helped to minimise isolation-induced transcription and likely reduced the sampling bias in singe-cell isolation that normally favours the sequestration of high-quality cells from less damaged, less fibrotic regions of the lesion (Denisenko et al., 2020; Haimon et al., 2018). Based on GO enrichment analysis, Zhou et al., 2020 found that the averaged myeloid cell expression pointed towards acute proliferation, in line with previous findings for microglia but not macrophages (Bellver-Landete et al., 2019; Greenhalgh & David, 2014; Milich et al., 2021; Noristani et al., 2017), and migration. At 7 dpi they observed migration and axon guidance, but not lipid-processing, indicating that not all myeloid cells perform this function at 7 dpi. By 14 dpi they observed that the averaged myeloid expression was enriched for GO terms regarding cell adhesion and extracellular matrix organisation (Zhou et al., 2020). The authors used this sequencing experiment to generate a hypothesis about the role of

microglial Plexin-B2 in axon guidance and regeneration post-SCI, which they then tested using several *in vitro* and *in vivo* contusion SCI model approaches. Ultimately, they demonstrated a role for myeloid Plexin-B2 in corralling infiltrating cells and the formation of the microglia scar (Zhou et al., 2020). Despite these exciting findings, this study was not without its caveats. Tamoxifen, which was administered just days before SCI supresses the T-cell immune response while in circulation, which would alter the course of the SCI pathology (Behjati & Frank, 2009). Additionally, there are concerns over the ability of nuclei-sequencing to accurately reflect the microglia transcriptome (Thrupp et al., 2020). Specifically, a comparison between 10x Genomics single nuclear RNAseq (snRNAseq) and scRNAseq of human microglia found that while snRNAseq was sufficient to identify the major neural cell types, it could not reproduce the activated microglial subpopulations observed in scRNAseq (Thrupp et al., 2020). Driving this discrepancy, the authors identified a depletion of about 1% of genes in snRNAseq data *vs* the scRNAseq data, and notably. These genes included several well-established players in microglial activation (APOE, SPP1,CD74, CST3). Despite these caveats, this bulk RNAseq INTACT study by Zhou et al., 2020 demonstrated that even in the single-cell era, carefully planned bulk RNAseq experiments are an inexpensive and effective tool for addressing novel biological questions.

scRNAseq Studies

An underlying theme within the bulk RNAseq literature is that the characterization of the SCI pathology with single-cell acuity will greatly contribute to our understanding of the complex pathology. As scRNAseq studies within the context of SCI exist have only emerged within the past 1.5 years, this aspires to be an exhaustive overview of the current peer-reviewed literature.

Two recent studies have employed scRNAseq using the 10X Chromium platform to better understand regenerative model organisms in the context of SCI. One study investigated the role of neural progenitor cells (NPCs) in regeneration after *Xenopus* tadpoles tail amputation (Kakebeen et al., 2020). These organisms undergo scar-free healing and complete spinal cord regeneration after these injuries and might identify targetable pathways to promote regeneration in the mammalian system. This study combined scRNAseq and Assay for Transposase-Accessible Chromatin (ATAC)-seq to find that NPCs in this regenerative system

undergo early neuronal differentiation followed by late phase proliferation. They also identified several transcription factors that they then demonstrated to be drivers of these functions, with notable parallels to neural crest development in mice (Kakebeen et al., 2020). Whether this study will lead to putative therapeutic options will, of course, require substantial further investigation in mammalian systems.

To investigate regeneration in a mouse model of SCI, Li et al., 2020 utilised neonatal mice after thoracic compression injuries; like *Xenopus* tadpoles, these animals undergo scar-free healing. The authors used enzymatic dissociation followed by FACS to isolate $CD45^{low}CD11b^{+}$ cells for scRNAseq (Y. Li et al., 2020). Their FACS protocol isolated not only the desired microglia but also other myeloid cells, lymphocytes, astrocytes, and oligodendrocytes. Thanks to the single-cell resolution of their technique, they were able to investigate the pure microglia populations. Interestingly, they found that neonatal microglia contribute to regeneration by secreting fibronectin and peptidase inhibitors and other anti-inflammatory molecules (Y. Li et al., 2020). They verified these findings through several *in vivo* and *ex vivo* experiments, including the transplantation of neonatal microglia into adult mouse models of thoracic compression injuries, which improved axonal regeneration, but no functional outcomes were reported. Nonetheless, this study used scRNAseq and complementary experiments to identify exciting mechanisms that can be further investigated to improve regeneration after SCI.

To investigate the spatial and transcriptional heterogeneity of mature oligodendrocytes after SCI, an extensive transcriptomics study was performed comprising 10X Chromium scRNAseq, single molecule fluorescence *in situ* hybridization (smFISH), and *in situ* sequencing (ISS), a technique capable of sequencing as many as 39 targets *in situ* (Floriddia et al., 2020). This project utilised a thoracic dorsal funiculi transection mouse model of SCI to create white matter damage with minimal animal suffering. This study was mainly descriptive but found that transcriptionally and spatially distinct mature oligodendrocyte populations had comparable transcriptional responses to SCI in the acute phase but developed distinct expression patterns in the chronic phase. Notably, the authors defined acute as 14 dpi and chronic as 3 and 5 months post-injury (Floriddia et al., 2020). This study represents a useful

resource that will undoubtedly serve as a foundation for future mechanistic studies into mature oligodendrocyte function and remyelination after SCI.

The INTACT bulk RNAseq study described in the previous section (Zhou et al., 2020), combined this 3, 7, and 14 dpi transection data with a 10X Chromium scRNAseq data from the whole CNS (Wahane et al., 2021). The CNS cells were dissociated enzymatically from 10 mm of tissue of uninjured controls and after mild (45 kilodyne) thoracic contusion SCI at 5 dpi. In this study, the authors used the averaged myeloid cell expression data from the INTACT RNAseq to define an injury-activated microglia/macrophage (IAM) profile which they noted, unsurprisingly, shared some commonalities but was largely different than the recently described "disease-associated microglia (DAM) profile" (Keren-Shaul et al., 2017). They found that at 5 dpi, microglia and oligodendrocyte precursor cells were proliferating, in line with previous findings (Bellver-Landete et al., 2019; Greenhalgh & David, 2014; Milich et al., 2021; Noristani et al., 2017; Wahane et al., 2021). They also observed activated microglia populations in the healthy control samples, suggesting their enzymatic dissociation protocol induced a technical bias in their data (Hammond et al., 2019; Li et al., 2019; Milich et al., 2021; Tay et al., 2017). Additionally, they repeated the scRNAseq after inhibiting histone deacetylase 3 (HDAC3), a gene product of interest from their bulk analysis and identified a long list of HDAC3-dependent SCI response genes, which were related to synaptogenesis or neuroinflammation in microglia and macrophages, respectively (Wahane et al., 2021). This report demonstrates the value of perturbation or treatment studies in scRNAseq, including in the context of SCI and the importance of designing protocols that minimise isolation-induced transcription in microglia.

A final study performed 10X Chromium scRNAseq on CNS resident and infiltrating cells from 8 mm of SCI tissue at 1, 3, and 7 dpi (Milich et al., 2021). They utilised a moderate thoracic contusion model of SCI and enzymatic dissociation. They also observed activated microglia in controls (Hammond et al., 2019; Li et al., 2019; Milich et al., 2021; Tay et al., 2017). Further discussion of the myeloid-specific findings of this study will be presented in great detail in *Chapter 5*. However, in this descriptive article, the authors characterised many other cell types after SCI, with a particular focus on their interactions. Milich et al., 2021 used CellPhoneDB (Vento-Tormo et al., 2018) to predict interactions between two cell types based

on the expression of ligand-receptor pairs. This analysis generated new hypotheses pertaining to the involvement of inter-cellular signalling in key processes such as gliogenesis, gliosis, and fibrosis. For example, at 1 dpi, they found that an increased expression of *Tie2* receptor and *Angpt2* ligand in endothelial cells and Tip cells, respectively. *Angpt2* has been shown to lead to the disruption of blood endothelium. Within the first week post-SCI, *Angpt2* expression in Tip cells was replaced by the blood endothelium-stabilising ligand, *Angpt1,* which was upregulated in astrocytes. Together this time-resolved data suggests a role for astrocytes in promoting blood vessel stabilisation after SCI. However, further experiments are required to validate these observations. Ultimately, this study highlighted the value in sampling all the cells within a system and once again outlined the importance of careful microglial isolation.

Notably lacking from the current SCI transcriptomics literature is spatial and human data. Understandably, both have been difficult to acquire. However, the rapid advancement of spatial high-throughput transcriptomics technologies will hopefully be reflected in the SCI literature soon. Some of these approaches, such as GeoMX Whole Transcriptome Atlas (WTA) (Roberts et al., 2021), can effectively sequence the entire transcriptome *in situ* and, excitingly, can be performed in FFPE tissue. Thus, we might soon see a spatially resolved, perhaps even time-resolved, human SCI cell atlas, which would undoubtably consolidate decades of human and animal research and generate novel, clinically relevant biological questions.

1.3.4 High-Throughput Single-Cell Transcriptomics in CNS Myeloid Cells

There exist many bulk studies of CNS myeloid cells, both resident and infiltrating, but the advent of single-cell technologies has demonstrated that myeloid cells previously combined in bulk studies, namely microglia, CAMs, and macrophages, represent transcriptionally distinct populations in development [§], homeostasis, and disease. Furthermore, in development and pathology these cells adapt drastically different states, even compared to cells of the same type (Masuda et al., 2020). Thus, here I will focus on highlighting the key findings from the past 5 years of high-throughput single-cell transcriptomics applied to CNS myeloid cells.

[§] scRNAseq has been monumental in deepening our understanding of the heterogenous states and roles of microglia during development. However, the details of these findings are beyond the scope of this dissertation and thus, will not be discussed further.

Homeostatic microglia have been extensively studied using scRNAseq. These studies, largely performed in mice, have generally supported the findings of decades of research by observing adult homeostatic gene signatures related to surveillance, phagocytosis, and synapse pruning (Matcovitch-Natan et al., 2016). They have also led to the generation of distinct homeostatic microglia profiles that separate them from CAMs and other CNS cell types, highlighting the caveats of many older studies that averaged the transcriptome or functional implications of these two populations, inspiring new tools for *in vivo* studies, and aiding the annotation of subsequent transcriptomics datasets (Masuda et al., 2020; Prinz et al., 2017). Additionally, despite the stark heterogeneity of microglia profiles in development and in response to pathology, scRNAseq studies in mice have concluded that adult homeostatic microglia are largely homogenous, irrespective of their location within the CNS (Masuda et al., 2020). Notably, scRNAseq and cytometry by time of flight (CyTOF) studies of human brain tissue have observed heterogeneity under homeostatic conditions, suggesting greater heterogeneity in human microglia (Masuda et al., 2020). Furthermore, a comparative study investigated the expression of the human *vs* mouse microglial sensome, genes involved in surveying the environment, and found that only about half of these genes are conserved across species (Abels et al., 2021), highlighting the need for human tissue validation of mouse transcriptome observations. Studies of CNS myeloid expression profiles in human *vs* mouse pathology have reached similar conclusions (Friedman et al., 2018). Most of these homeostatic CNS myeloid studies were performed as comparisons or controls for developmental or disease studies, but nonetheless, scRNAseq has deepened our understanding of homeostatic microglia.

Our understanding of myeloid ontogeny largely precedes the single-cell era but understanding this ontogeny has greatly influenced the experimental design and interpretation of scRNAseq studies. The prenatal ontogeny of microglia was first confirmed through fate-mapping techniques (Ginhoux et al., 2010) and was quickly replicated (Schulz et al., 2012). Shortly after, the transcription factors Irf8 and Pu.1 were shown to be required for the differentiation of microglia from erythromyeloid precursors (Kierdorf et al., 2013). Thus, before scRNAseq, it was clear that microglia and haematopoietic macrophages represent ontogenically distinct myeloid cell types (Kierdorf & Prinz, 2017; Prinz et al., 2017). However, an extensive project comprising scRNAseq, transgenic fate-mapping mouse lines, parabiosis-

generated blood chimera mice, and *in vivo* imaging was required to demonstrate that CAMs are long-lived and share the microglial prenatal origin, in contrast to monocyte-derived myeloid cells, such as macrophages (Goldmann et al., 2016). The scRNAseq data also demonstrated that the transcriptional profiles of homeostatic CAMs are similar but distinct from microglia.

Now that myeloid cell ontogeny has become less ambiguous, scRNAseq studies have used this knowledge to investigate the effect of myeloid ontogeny on their roles during homeostasis and CNS pathology. For example, Jordão et al., 2019 used an automated and miniaturised version of the CEL-Seq2 (Hashimshony et al., 2016) scRNAseq platform to investigate CAMs *vs* microglia (Jordão et al., 2019). Under homeostatic conditions, CAMs *vs* microglia were identified based on anatomical region, with different regions being processed for scRNAseq separately. During experimental autoimmune encephalomyelitis (EAE), they used a fate-mapping mouse line to separate resident CAMs and microglia from infiltrating myeloid cells (Jordão et al., 2019). Together with *in vivo* imaging and clonal analysis, the authors demonstrated stark heterogeneity within myeloid cell populations, particularly during disease. They also showed that even during disease CAMs, except for choroid plexus macrophages, maintain their populations through local proliferation, and that despite their expression of MHCII, these cells are not required for antigen presentation and EAE disease progression (Jordão et al., 2019). Ultimately, this project uncovered novel functions and dynamics of myeloid cells during disease, generated new biological questions surrounding EAE progression, and importantly, also serves as a useful resource for annotating future scRNAseq myeloid cell datasets.

scRNAseq has been used to address a longstanding body of conflicting evidence surrounding sexual dimorphism in microglia, particularly during development and disease (Masuda et al., 2020). Early scRNAseq studies observed very little difference, if any, between sexes (Hammond et al., 2019; Tay et al., 2018). However, several later studies have suggested there are indeed sex-specific roles in the microglia response to disease pathology (Han et al., 2021; Ochocka et al., 2021; Rahimian et al., 2019). This discrepancy might be attributed to the increased sensitivity of the scRNAseq methods over time.

High-throughput single-cell transcriptomics have undoubtably revolutionised our understanding of myeloid cell heterogeneity during disease. A substantial advancement was the elucidation of the molecular mechanisms driving the switch from homeostatic to activated microglia (Keren-Shaul et al., 2017). Specifically, the authors combined MARS-seq, animal models of neurodegeneration, tissue pathology including smFISH, epigenetic assays, and *Trem2* KO mice to describe a DAM profile that arises during AD, amyotrophic lateral sclerosis (ALS), and ageing. They described the DAM phenotype as progressing through two sequential phases, hereon referred to as DAM1 and DAM2. DAM1 was characterised by the upregulated expression of the TREM2 adaptors, *Apoe* and *Tyrobp*, and the downregulation of homeostatic microglia markers (Figure 1.8A) (Keren-Shaul et al., 2017). DAM2 phase was described as Trem2-dependent and was characterised by the upregulation of lysosomal, phagocytosis, and lipid metabolism genes. It has been described as a neuroprotective, phagocytic, phenotype in mouse models of AD, ALS, ageing, and later multiple sclerosis (MS) (Deczkowska et al., 2018; Keren-Shaul et al., 2017; Krasemann et al., 2017), and had interesting transcriptional parallels with early post-natal microglia (Hammond et al., 2019). Notably, they found that the DAM phenotype was conserved between mice and humans (Keren-Shaul et al., 2017), and was later suggested to be a disease-independent microglia response (Deczkowska et al., 2018).

Very shortly after the description of the DAM phenotype, another study using bulk RNAseq, several animal models of neurodegeneration, tissue pathology, and *Apoe* KO mice identified two major microglial states, a TGFβ-dependent, homeostatic, neuroprotective state (M0) and a disease-associated state (MGnD) (Krasemann et al., 2017). The authors provided gene signatures for these two states (Figure 1.8B), which intersected with the DAM phenotype but were not identical, and they proposed a molecular mechanism behind the microglia activation pattern (Krasemann et al., 2017). Building on the knowledge that TREM2 is a sensor of neurodegeneration-associated molecular patterns, including myelin debris and apoptotic neural cell bodies (Y. Wang et al., 2015), Krasemann et al., 2017 demonstrated that apoptotic neurons activated the TREM2-APOE pathway. They showed that this resulted in the suppression of the M0 state, as observed in the DAM1 response. Notably, unlike Keren-Shaul et al., 2017, Krasemann et al., 2017 hypothesised that MGnD can have negative effects on

chronic CNS disease pathology as these microglia lose the ability to suppress T-cell proliferation and are no longer able to protect neurons from cell death.

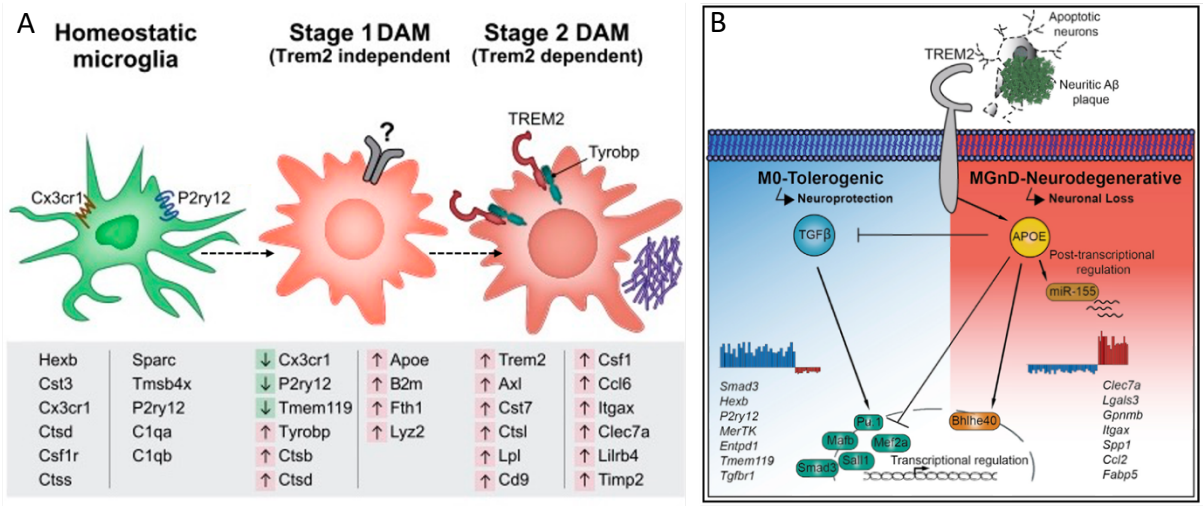


Figure 1.8 A) schematic of the transcriptional profile of homeostatic microglia, and the transcriptional changes associated with the transition to DAM1 and subsequently DAM2. The profiles were characterised in mouse models of ALS, ageing, and amyotrophic lateral sclerosis but have been hypothesised to be applicable to neurodegeneration in general. Figure adapted from (Keren-Shaul et al., 2017). B) A schematic of M0 *vs* MGnD microglia phenotypes, which have a common gene signature with the DAM phenotype but are not identical. Figure from (Krasemann et al., 2017)

Another study performed scRNAseq in mouse models of Alzheimer's disease (AD) using CEL-Seq2 but, despite using the same mouse model, the observed DAM profiles were not identical (Kubick et al., 2020; Tay et al., 2018). This highlights a few caveats of scRNAseq, which will be further discussed throughout this dissertation. To name a few: microglia are very sensitive to perturbations and so the isolation protocol can greatly impact the results; the sensitivity of the cell isolation platform and the design of the analysis workflow can also lead to slightly different conclusions.

As technology progresses, future *in situ* high-throughput transcriptomics experiments, particularly those with single cell resolution such as the much anticipated addition of single cell resolution to the 10x Genomics Visium platform, will likely challenge some of the scRNAseq observations, which may be affected by isolation-induced or sampling bias artefacts. Nonetheless, scRNAseq paired with complimentary approaches such as tissue pathology, fate-mapping, and functional studies, has greatly advanced our understanding of CNS myeloid cell biology in development, homeostasis, and disease. It has also highlighted an

interesting involvement of lipid metabolism in CNS pathology that is likely to be the target of many upcoming therapeutic approaches (Loving & Bruce, 2020).

1.4 Dissertation Outline

This dissertation is a summary of my PhD work in which I endeavoured to characterise the time-resolved transcriptional heterogeneity of CNS resident and infiltrating myeloid cells in response to SCI. To this end, I employed two transgenic mouse lines to collect myeloid cells from mouse models of SCI. I then integrated these datasets to create a time-resolved SCI myeloid cell atlas, which I analysed to generate a prospective map of myeloid cell dynamics across the SCI pathology. Ultimately, this project aimed to provide a better understanding of the role of myeloid cells during the secondary injury, with the hope that this fundamental knowledge might eventually lead to effective therapies to decrease the long-term disabilities of SCI patients.

Chapter 2 describes the methods I used to generate and maintain the thoracic contusion SCI mouse model and the sham-injured controls. In subsequent chapters, I utilise this model to generate the scRNAseq data and the tissue for the *in situ* validation of select findings.

Chapter 3 summarises the collection and pre-processing of the scRNAseq data from the first transgenic mouse line. This line labelled *Cx3cr1*$^{+}$ myeloid cells with a YFP reporter, enabling the efficient isolation of these cells via FACS. The isolation protocols and analysis workflow established in this chapter was largely conserved for the collection and analysis of the data from the second transgenic mouse line in *Chapter 4*.

Chapter 4 outlines the collection and pre-processing of scRNAseq data from the second transgenic mouse line and its integration with the previous dataset. This transgenic line aimed to differentially label infiltrating *vs* resident myeloid cells, allowing them to be isolated via FACS and thus unambiguously distinguished within the scRNAseq data. In this chapter I also describe how I have made this atlas publicly available and easily accessible for other researchers.

Chapter 5 describes the characterisation of the time-resolved SCI myeloid cell atlas, comparing it with previous findings and interpreting the dataset in its own right. Here I finish by summarising the myeloid cell atlas with a map of proposed phenotypes adopted across the SCI pathology.

Chapter 6 discusses the key caveats of the main technique utilised in this project, scRNAseq, considers the future directions of transcriptomics in SCI research, and provides a final overview of this dissertation.

Chapter 2 GENERATING THE MOUSE MODEL OF SPINAL CORD INJURY

"The more clearly we can focus our attention on the wonders and realities of the universe about us, the less taste we shall have for destruction."

—Rachel Carson

DISCLOSURE

The methods described in this chapter are expanded from the methods section of the bioRxiv pre-print: Hamel, R. *et al.* Time-resolved single-cell RNAseq profiling identifies a novel Fabp5-expressing subpopulation of inflammatory myeloid cells in chronic spinal cord injury. *bioRxiv*, 2020.2010.2021.346635, doi:10.1101/2020.10.21.346635 (2020). I wrote the manuscript for this pre-print in its entirety. Luca Peruzzotti-Jametti, Stefano Pluchino, and John Marioni edited the manuscript.

2.1 INTRODUCTION

SCI is a highly debilitating pathology afflicting thousands of individuals annually with no fully restorative treatment, resulting in substantial personal, societal and financial cost (Marinelli et al., 2019). To improve clinical outcomes, a better understanding of the SCI pathology is required. Unfortunately, studying the pathology in human tissue is limited by the scarcity of tissue samples, the inability to perform controlled, hypothesis-driven studies, such as loss of function experiments, and the commonly used FFPE preservation method, which presents challenges for molecular-level research. As such, researchers have turned to alternative means of investigating the SCI pathology.

Given the highly dynamic cellular and systemic responses involved in the SCI pathology, it is not currently possible to accurately model its full complexity *in vitro* or *in silico*. Thus, laboratory animals are still required for pre-clinical SCI research (Alizadeh et al., 2019; Sharif-Alhoseini et al., 2017). Mice are common models of SCI because their genetic, pathophysiological, and behavioural traits closely, albeit not perfectly (*1.2.2 Species*), recapitulate the human condition, providing opportunity to study these clinically relevant characteristics in a controlled setting. Furthermore, the commercial availability of transgenic mice allows for a vast number of gene-level applications, including observing and altering the expression of genes of interest, ultimately allowing the investigation of genetic programs contributing to the SCI pathology. In summary, to fully understand the dynamics of myeloid cells involved in secondary damages after SCI, animal models of SCI are currently indispensable.

The mouse model of contusion SCI aims to recapitulate the most common form of human SCI (Sharif-Alhoseini et al., 2017), and while cervical injuries are more common in humans, thoracic injuries are the most commonly studied (*1.2.1 Injuries*). Rodents, compared to the other species, allow the generation of relatively cost-effective, highly reproducible injuries and, like other small animal models, require rudimentary housing and husbandry, which is accessible at most research institutions (Lilley et al., 2020). Furthermore, methods to evaluate and compare behavioural and pathophysiological outcomes across studies and conditions have become well established for mouse models of SCI. However, while mouse models of

contusion SCI develop a secondary injury that resembles human contusion in many ways, fundamental differences exist in the size, anatomy, and pathophysiological response to the SCI, which can act as a barrier to translatability. Thus, researchers have also employed larger mammals, such as pigs and non-human primates, as models of SCI, but these are significantly constrained by both ethical and financial barriers (Alizadeh et al., 2019; Sharif-Alhoseini et al., 2017).

In this chapter, I describe the methods I used to generate and refine a mouse model of contusion, thoracic SCI. Then, I employ a well-established behavioural assessment to evaluate these models, ensuring their reproducibility, before ultimately employing these models to generate and begin to validate a time-resolved myeloid cell atlas after SCI.

2.2 METHODS

I performed all experimental animal procedures in accordance with the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB). The animal work was covered by the PPL 7008840 (to Stefano Pluchino).

2.2.1 MICE

Male and female mice aged 8 to 10 weeks and weighing 18-25g were bred and housed in a pathogen-free, temperature-controlled animal facility with 12hr/12hr light/dark cycles and *ad libitum* access to chow and water. I used the following strains for this project: Cx3Cr1^{CreERT2} strain name: *B6.129P2(Cg)-Cx3cr1$^{tm2.1(cre/ERT2)Litt/WganJ}$* to isolate *Cx3cr1*-YFP^{+} myeloid-lineage cells; TdTomato^{flox}, strain name: B6.Cg-Gt(ROSA)26Sor$^{tm9(CAG-tdTomato)Hze/J}$ crossed with *Cx3cr1*^{CreERT2} (*Crematо*) to distinguish between YFP^{+} infiltrating myeloid cells and double-positive RFP/YFP CNS resident microglia (MG), as recently described (Jordão et al., 2019). Both strains were purchased in 2018 from Jax. I also used C57BL/6 (WT) mice for tissue pathology and FACS, which were purchased from Charles River as needed.

2.2.2 PRE-OPERATIVE PROCEDURES

With the assistance of qualified University Biomedical Services technicians, I deeply anaesthetized the animals with isoflurane (4% induction, 2% maintenance) in oxygen (1.5 l/min), provided buprenorphine (Temgesic, RB Pharmaceuticals; pre- and post-operatively),

applied ocular ointment to prevent the eyes from drying, shaved the hair on the back of the mice, and swabbed it with a germicide.

2.2.3 LAMINECTOMY

I performed the laminectomy in line with the Pluchino lab's published protocol (Braga et al., 2020). First, I placed the anesthetised mice in the prone position and under a surgical microscope, created a dorsal midline incision over the thoracic vertebrae with a sterile scalpel and separated the paravertebral muscles using spring scissors (Fine Science Tools). I identified T12 as the apex of the dorsal aspect and performed the laminectomy using Dumont #2 laminectomy forceps (Fine Science Tools) and spring forceps (Fine Science Tools), leaving the dura intact. When the dura was punctured, there was an immediate reddening and swelling of the tissue. In this rare case, the animal was culled. To maintain vertebral column stability, I did not remove the lateral part of the vertebra at the site of laminectomy. The extension of the laminectomy was consistent between animals at approximately 3 mm in width and 5 mm in length, sufficient to allow room for the 1.3 mm diameter impactor tip.

2.2.4 INDUCTION OF CONTUSION SCI

For the SCI cohort, I induced a bilateral contusion injury on the exposed spinal cord at T12 using the Infinite Horizon (IH) impactor device (Precision Systems and Instrumentation, Lexington, KY; Figure 2.1) as previously described (Braga et al., 2020; Cusimano et al., 2012; Scheff & Roberts, 2009). The IH device is designed for the induction of a standard-force injury to the spinal cord of small rodents, and it allows the programming and monitoring of the desired force, the actual force, and spinal cord displacement. A limitation of this impactor is the IH Adson forceps, which are not fully adept at securing the spinal column and likely add a level of inconsistency to the parenchymal injury and functional outcomes (Cheriyan et al., 2014).

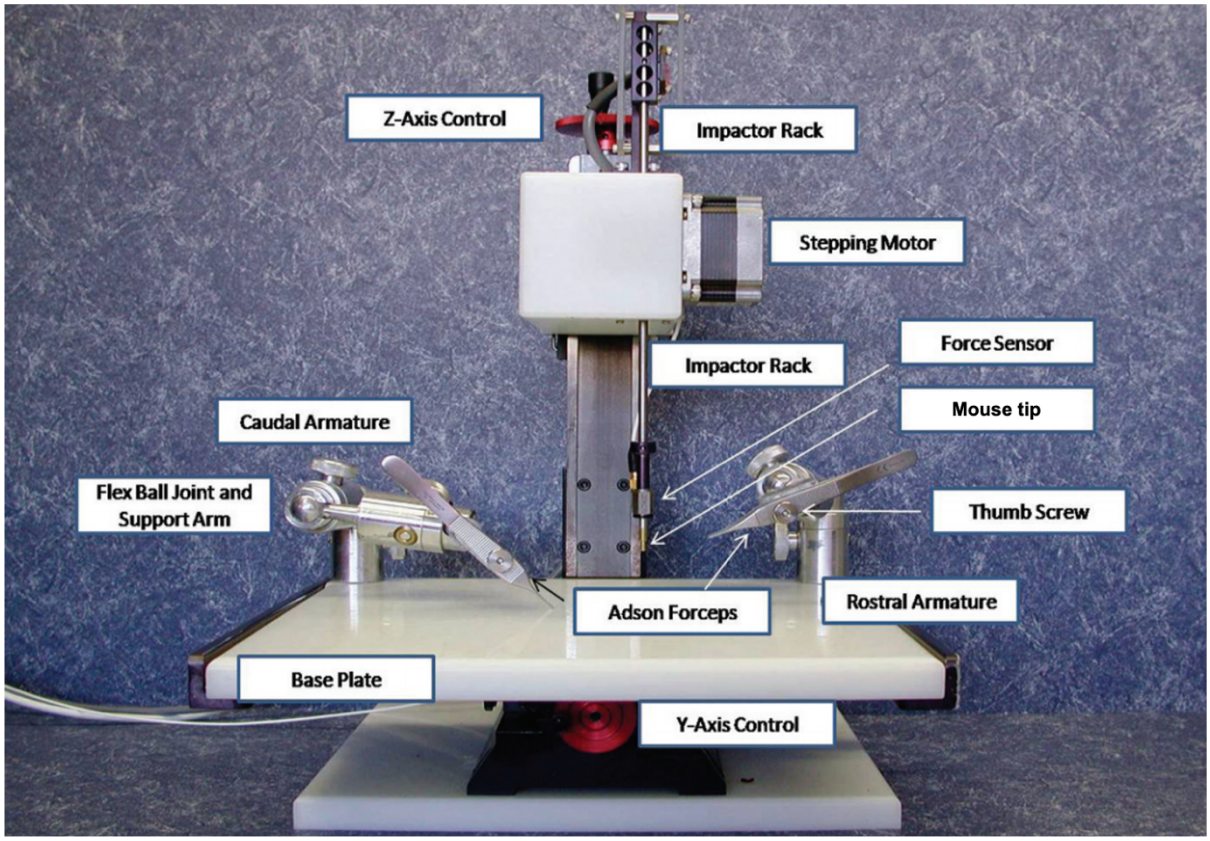


Figure 2.1 A front view of the IH impactor device. After the laminectomy, the mouse is placed in the prone position and the spinal column is secured using the Adson forceps. The mouse tip is aligned over the exposed cord using the y- and z-axis controls. The stepping motor drives the impactor rack into the exposed cord. During this process, the force sensor detects the point at which the desired force has been applied, at which point the impactor rack is retracted. The time, actual force, and displacement are recorded. Image adapted from (Scheff & Roberts, 2009).

To induce the injury, I aligned the mouse-impacting tip over the exposed cord and centred it over the central vein, avoiding any overlap with the transverse processes. Then, I induced a moderate contusion injury (70 kilodyne force) using the IH impactor and recorded the actual force and spinal cord displacement of the impact. For the SCI condition, I only included mice if the actual force was 70±5 kilodyne and the spinal cord displacement was 700±200 µm. For the sham-injured control group (Ctrl) I performed the laminectomy but omitted the induction of contusion SCI. After the injury, I closed the incision with 7-mm AutoClips (Fine Science Tool).

2.2.5 Post-Operative Care

After SCI induction or sham-injuries, mice recovered in a heated chamber. I performed post-operative care as described in the PPL 7008840. Briefly, I administered analgesics within 7 hours post-operation. For the first 7 dpi, a heat pad was secured to the outside bottom of the cage, and I checked the mice twice a day for signs of weight loss, urinary retention, pain,

infection, and re-opening of the incision. Additionally, I expelled the bladders of mice subjected to SCI by applying pressure to the bladder twice a day until urinary function was regained, or the endpoint was reached. After 7 dpi or post-laminectomy, I removed the AutoClips and continued to weigh and monitor the mice for the duration of the study plan.

2.2.6 Refinement

At the beginning of the study, in line with standard procedures at the University Biomedical Services animal facility and current refinement guidelines for mice post-SCI (Lilley et al., 2020), after surgery, the mice recovered from anaesthetics in a heated chamber overnight. Unfortunately, the rate at which mice reached their humane endpoint (due to adverse effects such as 20% weight loss, prolapse, and self-mutilation) before D10 was 47.22% (Figure 2.2). Note that mice were only included in this comparison after I was proficient in the surgery technique (> 1 year experience). When handling these mice, I observed that those with severe adverse effects warranting a humane endpoint were cold to touch, despite the heat pad. Thus, I began to house the mice in a heated chamber until their weight stabilised (3-5 dpi). A one-sided Fisher's exact test confirmed that keeping mice in a heated chamber until their weight stabilised decreased the number of mice that reached their humane endpoints before their scheduled time point ($p < 0.05$). Given this observation, I added this refinement to the lab's new PPL application.

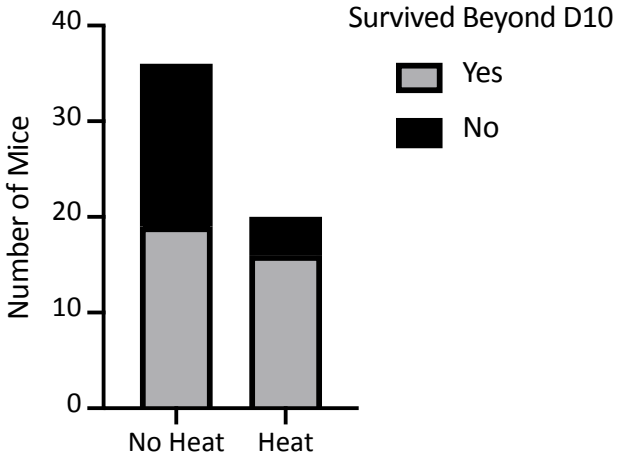


Figure 2.2 Bar plot depicting the number of mice surviving beyond 10 dpi with or without 3-5 days in a heated chamber. Recovering mice in a heated chamber substantially reduced the occurrence of adverse effects leading to humane endpoints.

2.3 Behavioural Assessment

I assessed hindlimb motor performance using the open-field Basso Mouse Scale (BMS) (Basso et al., 2006), with scores ranging from 0 (complete hindlimb paralysis) to 9 (healthy). I scored the animals for 4 minutes in an open field prior to the SCI induction and 1, 3, 5, 7, 10, 14, 18, and 21 dpi to confirm mice displayed the level of hind-limb locomotor impairment expected from a moderate contusion SCI. Only animals with left and right hindlimb BMS scores within 2 points of each other and a BMS score of 0 at 1 dpi were used. For statistical analysis of the BMS scores, I took the average scores of the left and right hind limbs resulting in a single BMS score for each animal. BMS scores changed as a function of time (Figure 2.3), as expected (Two-Way Mixed ANOVA; $F(7, 112) = 38.64$; $p < 0.0001$), however I did not observe behavioural differences between male and female mice over time (Two-Way Mixed ANOVA; $F(7, 112) = 0.5940$; $p = 0.76$).

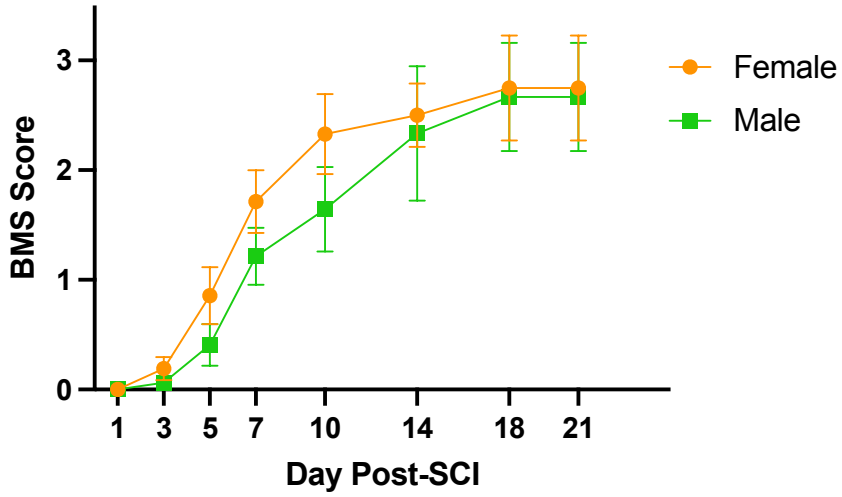


Figure 2.3 Basso Mouse Score (BMS) from male and female mice. Data are mean (± SEM) and have been collected from n≥4 mice per dpi.

2.4 Summary

Mouse models of contusion SCI are useful tools for furthering our understanding of the dynamic SCI pathology. The model is not without caveats, but advantages such as genetic modifications and manageable ethical concerns help to balance these out. The induction of contusion SCI is a well-established procedure both in the Pluchino lab and in the field of SCI research. While progress has been made to refine the technique and minimise adverse effects, future studies could look to further improve post-op care perhaps by extending the housing period in a heated chamber, although this preliminary observation requires

replication. A major refinement to the technique was the use of computerised impactor devices and standardised behavioural scales like the BMS, which allows for increased reproducibility across studies and ultimately reduce the number of animals needed for each study. As this PhD project investigated myeloid cells after SCI from both male and female mice, it was important to investigate sex-driven behavioural differences. The lack of such differences supported the use of male and female mice in this study and supports the effort to include female animals in pre-clinical studies.

Chapter 3 GENERATING THE TIME-RESOLVED MYELOID CELL ATLAS

"As always in life, people want a simple answer . . . and it's always wrong."

— Susan Greenfield

DISCLOSURE

All methods were planned, trialled, and performed by me, unless otherwise stated. Bryan Yu, a visiting undergraduate medical student from the University of Hong Kong, assisted me with many of the *ex vivo* procedures.

The methods described in this chapter are expanded from the methods section of the bioRxiv pre-print: Hamel, R. *et al.* Time-resolved single-cell RNAseq profiling identifies a novel Fabp5-expressing subpopulation of inflammatory myeloid cells in chronic spinal cord injury. *bioRxiv*, 2020.2010.2021.346635, doi:10.1101/2020.10.21.346635 (2020). I wrote the manuscript for this pre-print in its entirety. Luca Peruzzotti-Jametti, Stefano Pluchino, and John Marioni edited the manuscript.

3.1 INTRODUCTION

The ability of single-cell transcriptomics to uncover cellular-level heterogeneity, even in seemingly homogenous populations, has transformed our understanding of cellular biology. In the context of myeloid cells after SCI, scRNAseq has allowed us to deconvolute the dynamic and heterogenous response that occurs during the secondary injury at an unprecedented resolution. This high-throughput approach generates immense quantities of data with unique computational challenges; thus, the analysis of scRNAseq data must be planned and executed just as carefully as the sample collection.

Myeloid cells are of particular interest after SCI as they are major players in the inflammatory response in both humans and mice (Alizadeh et al., 2019; Gensel & Zhang, 2015). Furthermore, in contrast to peripheral lesions, SCI is characterised by persistent inflammation and the number of myeloid cells in the spinal cord post-SCI remain elevated well into the chronic phase (Donnelly & Popovich, 2008) and likely perpetuate the inflammation that characterises the chronic wound state. However, over the course of the SCI pathology, myeloid cells play complex and dynamic roles with both beneficial and detrimental effects (Gaudet & Fonken, 2018) (*1.1.3 A Closer Look at the Immune Response to SCI*). Thus, a time resolved analysis at single cell resolution would help to unravel the heterogenous roles of myeloid cells post-SCI.

Here, I developed a single cell isolation protocol to extract *Cx3cr1*$^{+}$ myeloid cells from transgenic mouse models of contusion SCI across the acute, subacute, and early chronic phases of SCI for scRNAseq. Significant effort was made to minimise extraction-induced transcription. Other researchers have confirmed that microglia are highly sensitive to perturbations and have shown that the isolation of these cells for scRNAseq can produce activated clusters if extra precautions are not taken (Haimon et al., 2018; Marsh et al., 2020). Using this *Cx3cr1* dataset, I also established a pre-processing, quality control and preliminary analysis workflow based on published approaches, most of which have recently been summarised in an online book (Amezquita et al., 2020). The code I wrote for handling the data in this chapter is publicly available on my GitHub (https://github.com/regan-hamel/PhD_Dissertation).

3.2 Myeloid Cell Isolation

The most common method of isolating cells of interest for the generation of omics data is via FACS, which requires the cells to be fluorescently labelled. One approach to fluorescently labelling cells is to stain cell surface protein(s) that serve as cell markers (such as CD45/CD11b for myeloid cells) with fluorophore-conjugated antibodies. However, this approach has several caveats, especially for myeloid cells. Firstly, it requires staining and washing steps, which adds time to the protocol and is typically performed at room temperature (Li et al., 2019) or on ice with longer incubation periods (Hammond et al., 2019). Secondly, given the vast quantity of cells in suspension and the time constraint, antibody staining is inefficient compared to transgenic fluorescent labelling. When attempting to isolate optimal cell quantities for scRNAseq (~20,000) from small quantities of tissue (5 mm of the mouse spinal cord) and minimizing the time *ex vivo*, transgenic labelling is the superior approach.

In this chapter, I used a transgenic mouse strain to label and ultimately isolate myeloid-lineage cells for scRNAseq. The mouse strain was *B6.129P2(Cg)-Cx3cr1*$^{tm2.1(cre/ERT2)Litt/WganJ}$ (Parkhurst et al., 2013), purchased in 2018 from Jax, hereon referred to as *Cx3cr1*. This strain has constitutive expression of the enhanced yellow fluorescent protein (YFP) transgene and a tamoxifen-dependent Cre recombinase under the *Cx3cr1* promoter, ultimately labelling all *Cx3cr1*-expressing cells as YFP$^+$ (see *4.2*). With this mouse strain, I generate and maintained mouse models of contusion SCI or laminectomy-only controls as described in *Chapter 2*. At 1, 2, 3, 10, or 21 dpi, I deeply anaesthetised the mice with an intraperitoneal injection of 100 µl of pentobarbital sodium before performing transcardial perfusion with carbogen-bubbled, ice-cold, artificial cerebral spinal fluid (aCSF) for 7 minutes or until the liver was clear of blood. I chose to use aCSF (Tasic et al., 2016; Zeisel et al., 2018) to maintain the tissue in physiological-like conditions for as long as possible, minimising isolation-induced transcription. Furthermore, from the point of perfusion, I kept all tissue or cells at 4°C or on ice, to slow metabolic activities and prevent molecular changes associated with the isolation protocol.

After perfusion, I extracted the spinal columns and placed them in conical tubes of ice-cold aCSF. I voided the spinal column using the hydraulic extrusion method (Au - Richner et al.,

2017) using a 5 ml syringe filled with ice-cold aCSF. I cut a 5 mm section of the spinal cord, centred on the lesion. Then, I mechanically homogenised the tissue sections using a 7 ml Dounce Tissue Grinder (DWK Life Sciences) filled with 6 ml of homogenisation buffer that I concocted as follows: aCSF with 10 mM HEPES (Sigma) to maintain stable pH despite CO_2 released from the cells; 1% BSA (Sigma) to reduce cell clumping; 1 mM EDTA (Thermo Scientific) to chelate Ca^{++} and Mg^{++}, reducing cell adhesion; 10 mg/ml of DNAse (Roche) 3000U to degrade free-floating DNA and reduce cell clumping; and 40 units/µl of RNAse inhibitor (Invitrogen) to minimize the degradation of mRNA transcripts. I chose mechanical tissue dissociation over enzymatic dissociation because it is much faster (1 minute *vs* 30 minutes), can be performed on ice *vs* 37°C, and is less likely to cause conformational changes of cell surface receptors, which could elicit transcriptional changes.

After tissue dissociation, all pipetting was performed using wide bore pipette tips (Alpha Labs) to reduce shear strain and ultimately increase cell viability and quality. I filtered the dissociated tissue suspension through a 40 µm strainer. I pre-wet the strainer with aCSF to prevent cells from adhering to the dry mesh and I rinsed the homogeniser with 2 ml of homogenisation buffer to increase cell yield from these small tissue sections. Then, I removed myelin and debris from the cell suspension by adding 2.7 ml of isotonic 9:1 Percoll (Sigma) to 10X PBS. I gently mixed the samples by inverting them 10 times and then centrifuged at 800 g for 20 minutes at 4°C. The brake speed was set to 0 to preserve the myelin debris ring, which visibly layered at the surface. After removing the myelin debris ring, I washed away the Percoll by adding ~5 ml of ice-cold buffer of 5% autoMACS Rinsing Solution (Miltenyi Biotec) in 1X MACS BSA solution (Miltenyi Biotec) and then centrifuging the samples at 800 g for 5 minutes at 4°C.

I resuspended the pelleted cells in 200 µl of FACS buffer, which comprised Cell Staining Buffer (Biolegend) with 10 mM HEPES (Sigma), 1% BSA (Sigma), 1 mM EDTA (Thermo Scientific), 10 mg/ml of DNAse (Roche) 3000U, 40 units/µl of RNAse inhibitor (Invitrogen), and 7-AAD live/dead stain (Invitrogen) at a concentration of 1:50. Finally, I worked with the technicians at the Cambridge Phenotyping Hub to isolate live myeloid cells ($7\text{-}AAD^{-}$ YFP^{+}) using a BD FACS Aria III cell sorter. We set the sorter to 3-way purity and 20 psi with a 100 µm nozzle – these parameters allow for high accuracy sorting (minimising the isolation of non-myeloid [YFP^{-}] cell

types) with less pressure and shear stress applied to the cells compared to smaller nozzles and higher pressure. As the panel consisted of only YFP and 7-AAD, fluorochrome compensation was not required. We set the gates based on an unstained WT sample and a 7-AAD stained WT sample (Figure 3.1A). To increase cell yield but minimise the isolation of non-myeloid cells, we set the cell and singlet gates generously, but were strict on the live and YFP^{+} gates (Figure 3.1B).

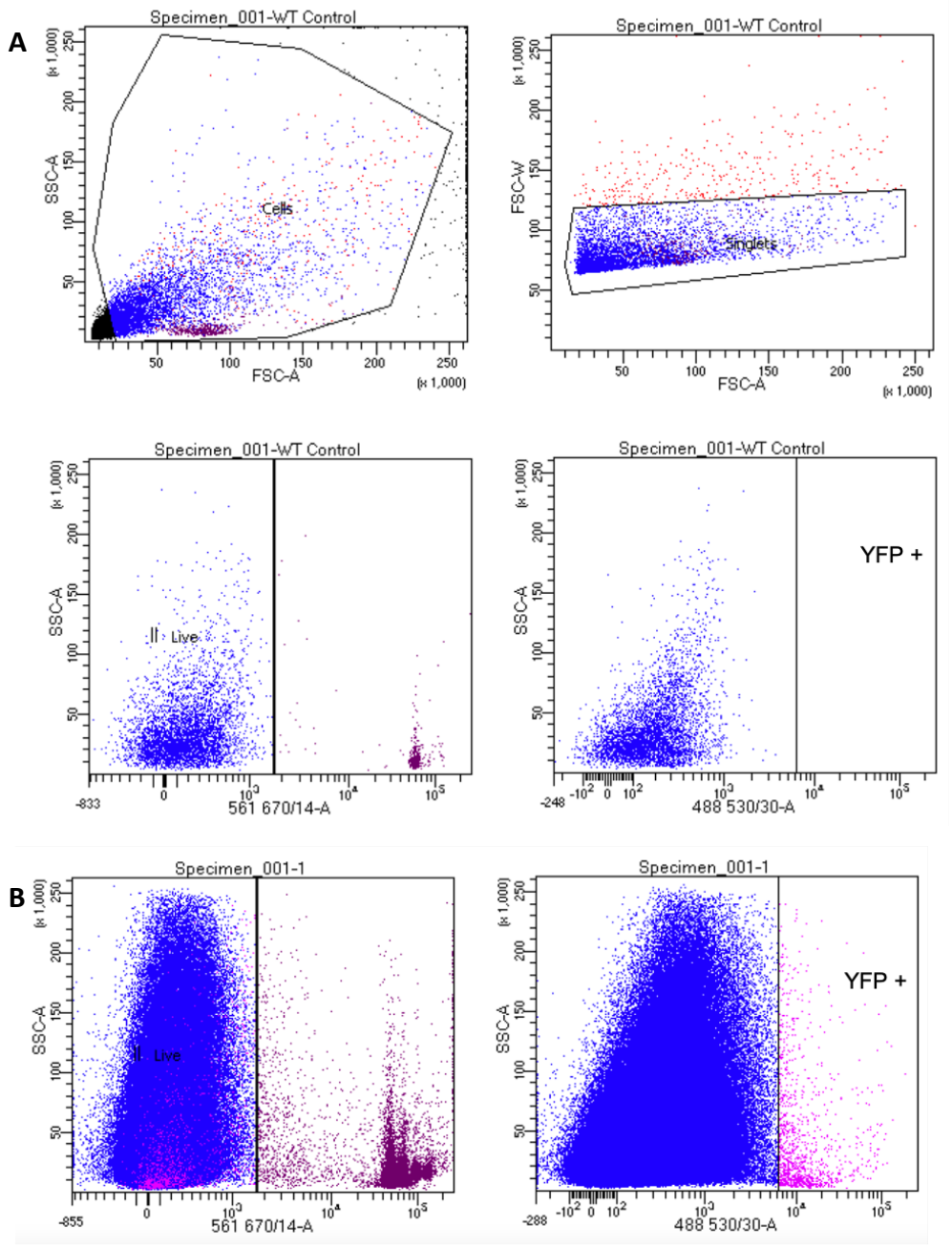


Figure 3.1 A representative FACS gating strategy based on a 7-AAD stained WT sample (A) for a 21 dpi *Cx3cr1* sample (B). SSC = side scatter; FSC = forward scatter

After scRNAseq and compiling the myeloid dataset as described below, I investigated the expression of FACS and dissociation-associated genes (van den Brink et al., 2017) in the Ctrl samples of this dataset, and compared it with Ctrl microglia from other scRNAseq studies (Arneson et al., 2018; Hammond et al., 2019; Van Hove et al., 2019; Zeisel et al., 2018). I found that, after applying the same QC thresholds to each sample, Ctrl microglia from this study expressed at least slightly lower levels of FACS and dissociation-associated genes, even compared to a sample that was processed without FACS (Figure 3.2).

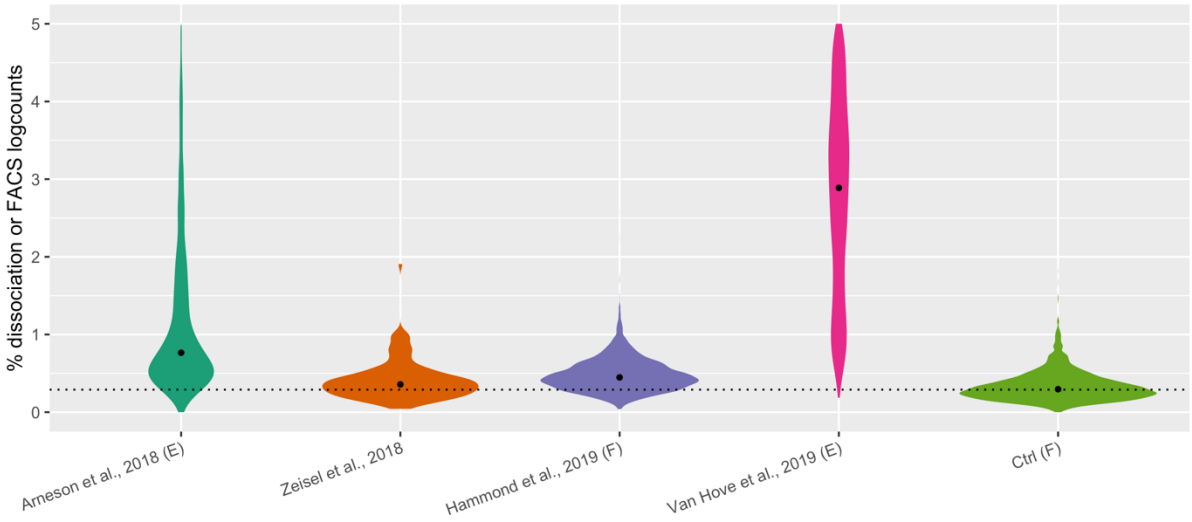


Figure 3.2 Violin plots showing the percentage of all logcounts originating from dissociation- or FACS-associated genes, per Ctrl microglia cell, across several droplet- based scRNAseq studies after applying the same QC metrics. Black dots represent the median. F= FACS, E= enzymatic dissociation (*vs*. mechanical). The dotted line represents the median for this study, which was lower than the others, even Zeisel et al., who did not use FACS (or any alternative method) to isolate cells.

3.3 Single-Cell RNA Sequencing

The staff at the CRUK Cambridge Institute Genomics Core sequenced the cells with single-cell acuity using the microdroplet based-platform, 10X Genomics Chromium Single Cell 3' Solution followed by Illumina sequencing.

3.3.1 10X Genomics Chromium Barcoding & Library construction

The 10X Genomics Chromium platform exploits the polarity differences between water and oil to encapsulate single cells, enzymes and nucleotides for reverse transcription, and gel beads into water droplets, cumulatively referred to as a gel bead in emulsion (GEM) (Figure 3.3A). For approximately 1% per 1000 cells, two cells are captured within a single GEM, which is known as a doublet. After GEM formation, the cells are lysed and the mRNAs are reversed transcribed into cDNA within the GEM, incorporating the unique barcodes into their

sequence. The templates for the barcodes are DNA oligonucleotide molecules of known sequences, which coat the gel beads (Figure 3.3B) and contain 4 segments: a read 1 primer that allows the barcode to be synthesized and read during sequencing; a 16bp cell barcode that is unique to each bead allowing us to demultiplex the samples, i.e., trace the mRNA back to the cell of origin after sequencing; a UMI of 10bp (version 2 chemistry) or 12bp (version 3 chemistry) unique to each oligonucleotide, allowing amplified transcripts to be deduplicated, i.e. traced back to their original mRNA, enabling us to discriminate between transcripts that were highly expressed *in vivo* versus those that were highly amplified *ex vivo*; and a poly(dT) tail allowing the barcoded oligonucleotide to bind the poly(A) tail of the mRNA. To ensure that the poly(A) anneals to the beginning of the poly(T), the poly(dT) primer also includes a VN anchor (where N is any nucleotide and V is not T).

Once the cDNA is barcoded by its cell and transcript of origin, the oil can be removed, and library preparation proceeds similar to bulk sequencing (Figure 3.3C). First, the cDNA is amplified, then enzymatically fragmented and size selected using SPRI beads. The fragments of optimal length (400-600bp) are then ligated to an 8bp sample index barcode that denotes the sample, a read 2 primer that allows the transcript to be synthesized and thus sequenced, and P5/P7 primers used for amplification on the Illumina flow cell. Finally, the full cDNA construct is amplified through PCR and then sequenced.

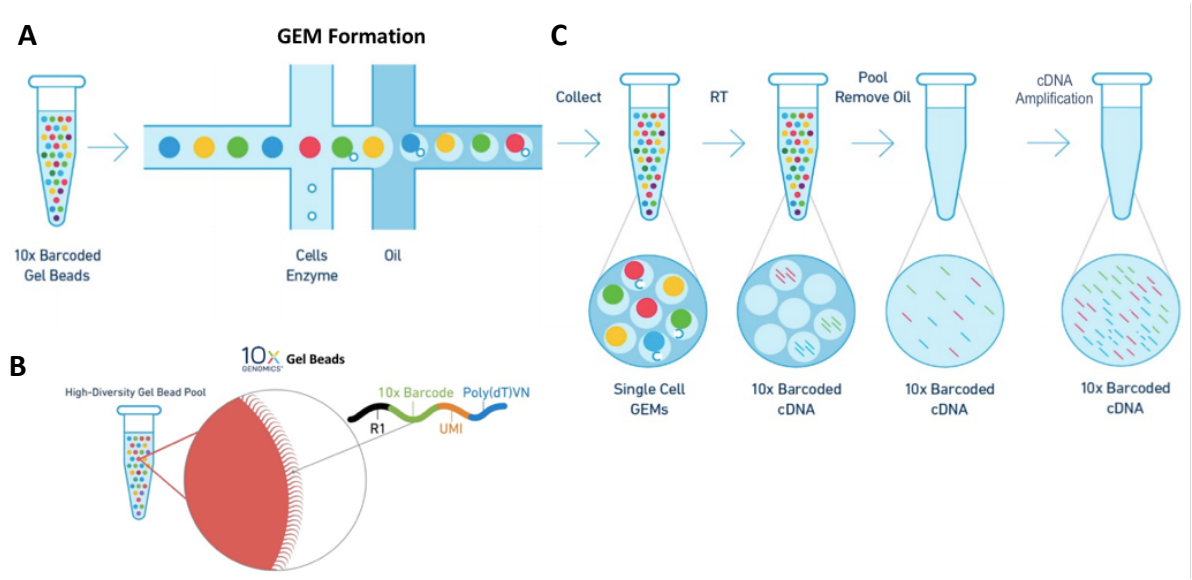


Figure 3.3 10X Genomics Chromium Single Cell 3' Solution workflow: A) GEM formation, B) Oligonucleotide coated gel beads, C) the library preparation for Illumina sequencing (Durruthy, 2019)

3.3.2 Illumina Sequencing

The cDNA libraries generated in January and March 2019 were sequenced on 2 lanes at a depth of at least 36,000 reads per cell on the HiSeq4000 (Illumina). All later samples were sequenced on the NovaSeq 6000 at a depth of at least 50,000 reads per cell. This change was made at the advice of the Genomics Core staff who determined that it improved base quality scores by 15-20%.

During sequencing, the 10X primer (read 1) directs the 26bp sequencing of the 10X barcode and UMI for each cDNA fragment, avoiding the unnecessary sequencing of the poly(A). The read 2 primer directs the sequencing of the transcript. In this project, the purpose of sequencing was to identify and quantify mRNAs, so it was not necessary to sequence entire transcripts. Thus, 98 (version 2) or 91 (version 3) cycles of sequencing were requested from the read 2 primer, as this was the optimal length recommended by 10X Genomics for effective alignment to the mouse transcriptome.

The output of the sequencer was a binary base call file, which registers the identity, location, and confidence level for each detected fluorophore on the flow cell in real time during sequencing. This file was converted to a FASTQ format via bcl2fastq conversion software (Illumina), and a preliminary quality checks were performed using FastQC (Babraham Bioinformatics, CRUK Cambridge Institute), a java-based software that quickly verifies the sequencing quality and can identify the source of the low-quality data if applicable.

3.4 Compiling the Time-Resolved Myeloid Cell Atlas

The vast quantity of data generated through scRNAseq requires careful pre-processing before it can be interpreted in a biologically meaningful way. The key steps in this workflow are outlined in Figure 3.4, each of which will be discussed in detail below. Each sample is barcoded separately but multiplexed with other samples of the same or similar collection date. From there, each sample is pre-processed in parallel, until the normalisation step. At this point, the samples are integrated and are then ready for further analysis and interpretation.

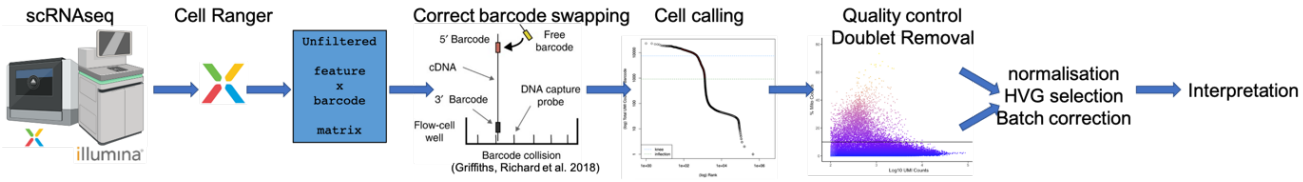


Figure 3.4 The general workflow for the compiling the myeloid cell atlas

3.4.1 Aligning the sequences to the mouse transcriptome

After obtaining the demultiplexed FASTQ files from the Genomics Core, I aligned the reads to the GRCm38 (mm10) mouse genome using 10X Genomics' *Cell Ranger* software (version 3.1.0). Notably, the gene *Clec7a* was not present in the mm10 reference genome supplied by *Cell Ranger*, so with the help of Dr Andrian Yang, a post-doctoral researcher in the Marioni group, I generated an edited mm10 reference genome to include this gene. The output from the *Cell Ranger* count pipeline includes an unfiltered feature-barcode matrix, which contains all the cell barcodes with at least one associated read.

I performed all remaining scRNAseq data analysis in R version 3.6.3, largely handling the data as a *SingleCellExperiment* object from the *SingleCellExperiment* package (version 1.8.0). All the packages I used are publicly available on the Comprehensive R Archive Network (https://cran.r-project.org), Bioconductor project (http://bioconductor.org), or their respective GitHub pages. The complete R analysis workflow used in this chapter is available on GitHub: https://github.com/regan-hamel/PhD_Dissertation.

3.4.2 Correcting Barcode Swapping

The two Illumina sequencers used in this study (HiSeq4000 and NovaSeq 6000) utilise patterned flow cells to increase cluster density compared to bridge amplification, ultimately reducing the cost and increasing the read quality (Sinha et al., 2017). However, this sequencing method has been associated with an effect called barcode swapping, in which multiplexed cDNA fragments are labelled with an inappropriate sample index barcode during the cluster-generating amplification step on the patterned flow cell. Barcode swapping might affect 2.5%-10% of reads (Griffiths, Richard, et al., 2018; Sinha et al., 2017) and has been shown to produce artefactual libraries that can be mistaken for cell libraries in droplet-based scRNAseq studies such as this one. To mitigate this effect, an algorithm has been developed to identify and remove reads from swapped sample index barcodes (Griffiths, Richard, et al.,

2018). Briefly, this approach assumes that two barcoded cDNA molecules from different samples are highly unlikely to have the same combination of UMI sequence, sample index barcode, and gene identity. Thus, when examining multiplexed samples, the algorithm identifies these "unlikely" molecules and removes them, unless they originate from the dominant sample.

Before further processing of the feature-barcode matrix, I used the aforementioned algorithm to remove swapped molecules by calling *swappedDrops* from the *DropletUtils* package (version 1.6.1) on the samples of a given sequencing batch. Three samples (D10, G11, G9) required a second round of sequencing to increase their depth, and in these cases, I was not able to correct for barcode swapping.

3.4.3 Cell Calling

To minimise the occurrence of doublets, the droplet-based single-cell isolation system described in Section 3.3.1 results in far more empty droplets than cell capturing-droplets. The identity of the latter is not known *a priori,* and empty droplets are not truly empty, but rather contain ambient transcripts released from damaged cells during the isolation process. Thus, there is a need for computational approaches to distinguish empty *vs*. cell-containing droplets.

Cell Ranger is the default approach for cell calling in 10X Genomics scRNAseq samples. Before *Cell Ranger* version 3, the pipeline called cells by assuming that cell-containing droplets corresponded to increased UMI counts (Lun et al., 2019). For a heterogenous sample, such as myeloid cells isolated *ex vivo* after SCI, *Cell Ranger's* original barcode filtering method was not the most appropriate approach.

Before *Cell Ranger* version 3 became available, I called cells from each sample's unfiltered feature-barcode matrix using the *emptyDrops* function from the *DropletUtils* package (Lun et al., 2019). This approach prevents cells with low-UMI counts from being discarded as empty droplets by profiling the background mRNA barcoded in empty droplets and then classifying cell-containing droplets as those that significantly differ from the background. To visually inspect the outputs of this approach, we can take the negative log probability of obtaining a

given cell barcode's profile from the background profile and plot it against the cell library size (UMI count) and map the called cells *vs* empty droplets (Figure 3.5). The *emptyDrops* filtered feature-barcode matrix for each sample was then used for all downstream processing.

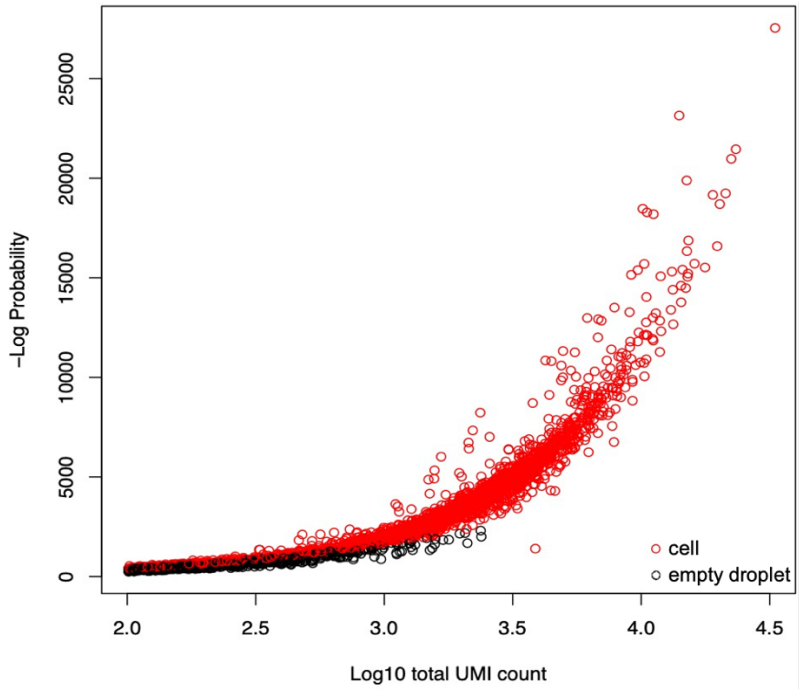


Figure 3.5 A diagnostic plot of the *EmptyDrops* output for sample G11. Each circle represents a cell barcode. As expected, called cells (red) have a higher -log probability or a much larger UMI count than empty droplets (black). Cells with low UMI counts but a relatively larger -log probability resulted in a FDR < 0.01 and were not discarded as empty droplets on the basis that their mRNA profiles significantly differed from the profiles of the null model.

A similar approach to *emptyDrops* was adopted in *Cell Ranger* version 3. I chose to continue using *emptyDrops* to allow for inspection into the cell calling statistics, however using the filtered feature-barcode matrices from *Cell Ranger* version 3 would have also been sufficient.

3.4.4 Quality Control

Before the gene expression of the feature-barcode matrices can be biologically interpreted, low quality cells must be removed; there are several standard approaches to accomplish this, with fixed *vs* adaptive thresholds being the most common (Amezquita et al., 2020). To determine which approach was the most appropriate for this dataset, I first used the *perCellQCMetrics* function from the *scater* package (version 1.14.6) (McCarthy et al., 2017) to investigate the cell-based QC metrics for each sample.

Within each sample, the distribution of library sizes[**] was similar to that of the complexity and in many cases, cells with lower complexity and smaller library sizes had higher proportions of mitochondrial transcripts (Figure 3.6). Based on these three metrics, most cell libraries in sample H9 were low quality, an observation supported by the fact that sample also had a low fraction of reads in cells (only 45.1%). Thus, sample H9 was excluded from further analysis.

Samples from the 10-Jan-19 collection date (F3, E3, D3), processed with version 2 chemistry, had smaller library sizes and fewer detected genes compared to the version 3 samples. Yet, they did not have the high proportion of mitochondrial transcripts observed in cells from version 3 samples with similar library sizes, suggesting that it was a difference in chemistry and not necessarily cell quality that led to reduced library size.

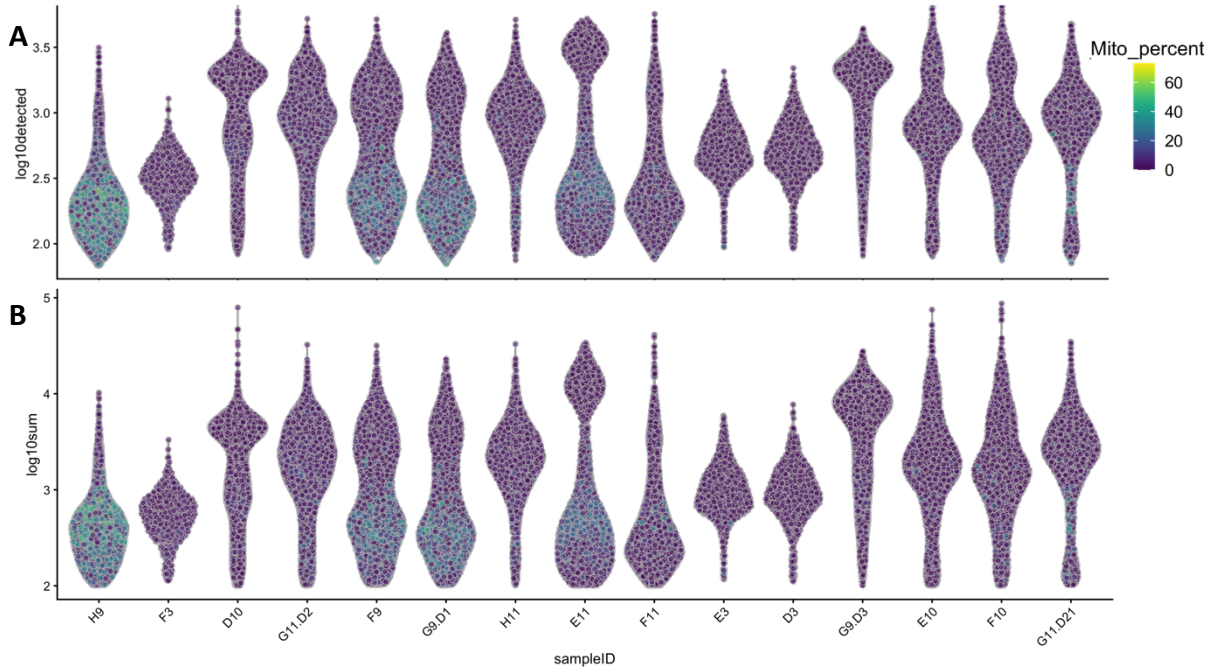


Figure 3.6 Violin plots depicting the log10 number of detected genes per cell (A) or UMI counts per cell (B) for each sample. Each cell is coloured by its proportion of mitochondrial transcripts. Cells with fewer detected genes and fewer UMI counts often expressed higher proportions of mitochondrial transcripts, suggesting these cells are of low quality.

Given that the *Cx3cr1* dataset comprised five different collection dates, I wondered whether the observed variation in QC metrics across the samples might correspond to the batch in

[**] Library size refers to the number of UMI counts per cell. Complexity refers to the number of detected genes.

which the samples were isolated and barcoded. This was generally true for samples of the same dpi collected on the same date (Figure 3.7).

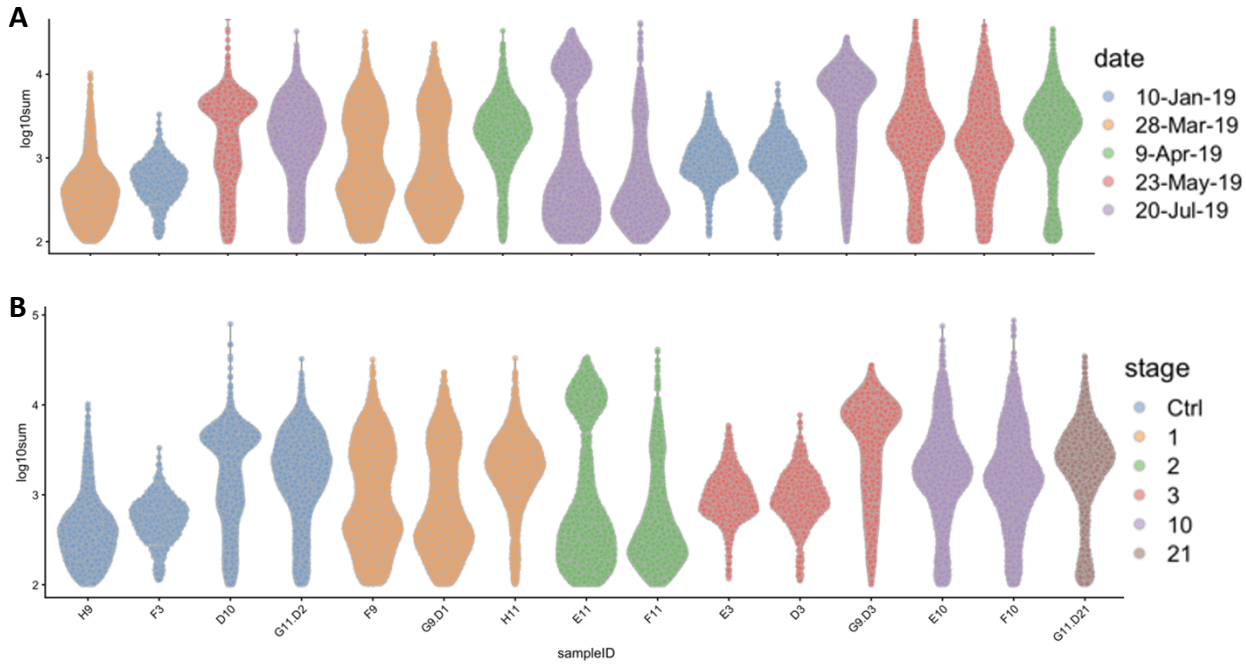


Figure 3.7 Violin plots depicting the log10 UMI counts per cell for each sample. Each cell is coloured by the collection date (A) or the dpi (B). Samples from the same condition that were collected on the same day appear to have similar distributions of cell library sizes.

After investigating the QC metrics, I utilized them to filter out low-quality cells. Unfortunately, since I isolated cells from the damaged epicentre of the spinal cord and many samples showed the hallmark signs of poor quality (small library, low complexity, and high proportions of mitochondrial transcripts; Figure 3.6), I could not assume that all samples showed typical features of high quality libraries and thus could not identify low quality cells with adaptive QC thresholds. Instead, I removed low-quality cells using fixed thresholds. For the proportion of mitochondrial transcripts, I chose to use a fixed threshold of 5% (Figure 3.8), in line with a recent study of mouse mitochondrial transcript proportions in high *vs* low quality cells in scRNAseq datasets (Osorio & Cai, 2021). I considered that this approach might remove metabolically active cells, however if such cells were present in the dataset, I would expect them to have larger library sizes, i.e., populate the upper right quadrant of the plot in Figure 3.8, which was not the case.

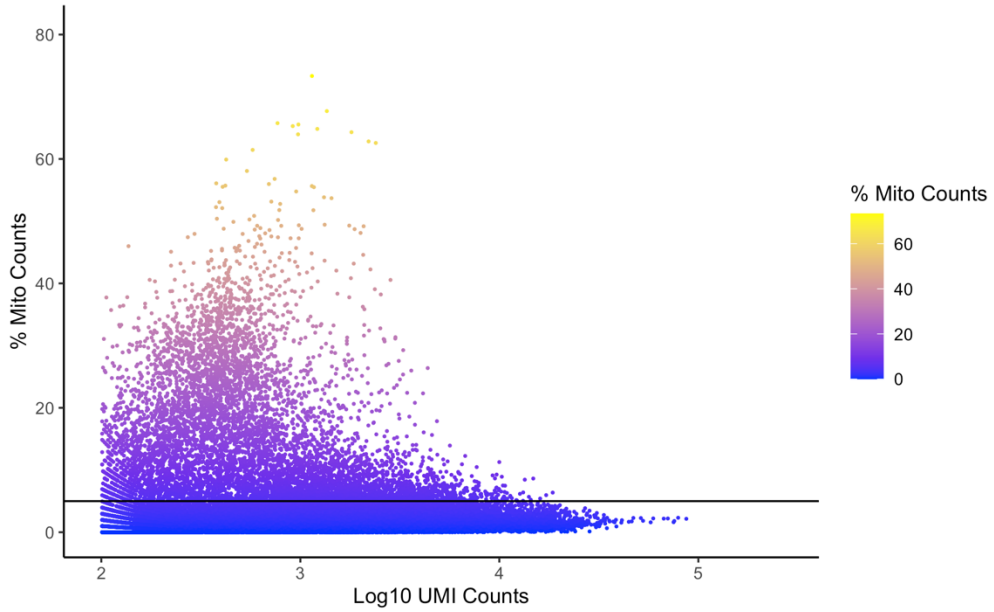
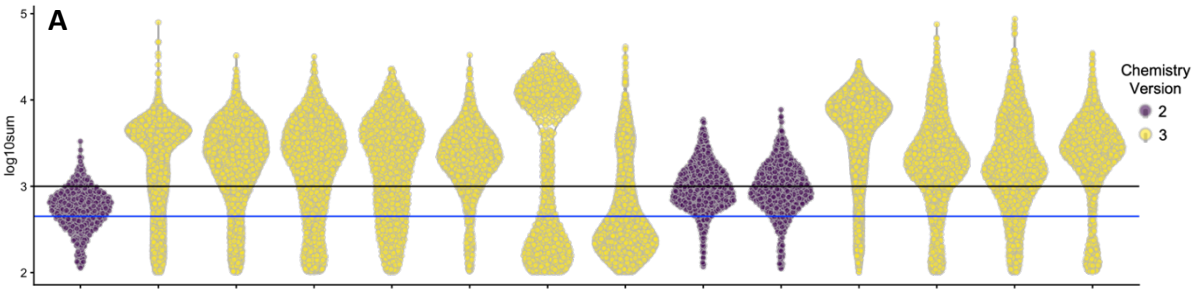


Figure 3.8 A plot of the log 10 library size vs the proportion of mitochondrial transcripts for all the Cx3cr1 samples except H9 (removed for QC). Each cell is coloured by its proportion of mitochondrial transcripts. Cells with more than 5% mitochondrial counts were removed as low-quality.

To further filter the *Cx3cr1* dataset, I set thresholds for the minimum number of UMIs and genes, which differed between version 3 and version 2 samples (Figure 3.9). These values were determined empirically but were similar to reported values for other myeloid cell scRNAseq datasets (Hammond et al., 2019; Zeisel et al., 2018). After applying these thresholds, sample F3 contained fewer than 100 cells and was removed as a failed sample. In total, 12,235 cell libraries passed the QC filtering.



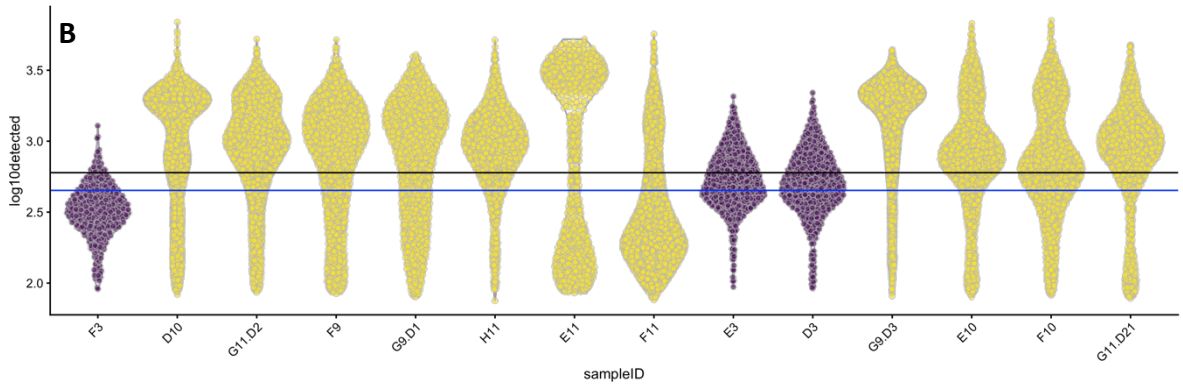


Figure 3.9 A violin plot of the *Cx3cr1* dataset after cells have been filtered by mitochondrial transcript proportions. Each cell is coloured by chemistry version. A) the log10 UMI counts per sample. The black line depicts the version 3 threshold of 1000 UMIs per cell; the blue line is the version 2 threshold of 450 UMIs per cell. B) The log10 detected genes per sample. The black line depicts the version 3 threshold of 600 genes per cell; the blue line is the version 2 threshold of 450 genes per cell.

Given the expected heterogeneity within the *Cx3cr1* dataset, I questioned whether my QC thresholds had erroneously removed certain myeloid cell types. To address this, I calculated the average counts per gene for the kept *vs* discarded libraries, then used these to calculate the average $\log_2$ fold change ($\log_2$FC) of each gene, as previously described (Amezquita et al., 2020). I visualised these results by plotting the average UMI count of a given gene *vs* its $\log_2$FC and observed an increase in certain genes in the discarded libraries. I investigated the genes with a $\log_2$FC greater than 0.25, and as expected, many were mitochondrial genes. However, based on PanglaoDB (Franzén et al., 2019), many of these genes were expressed in non-myeloid CNS or endothelial cells (Figure 3.10A). Conversely, the genes with the highest increase in the retained libraries were myeloid cell genes (Figure 3.10B), suggesting that the QC thresholds removed the appropriate cell barcodes.

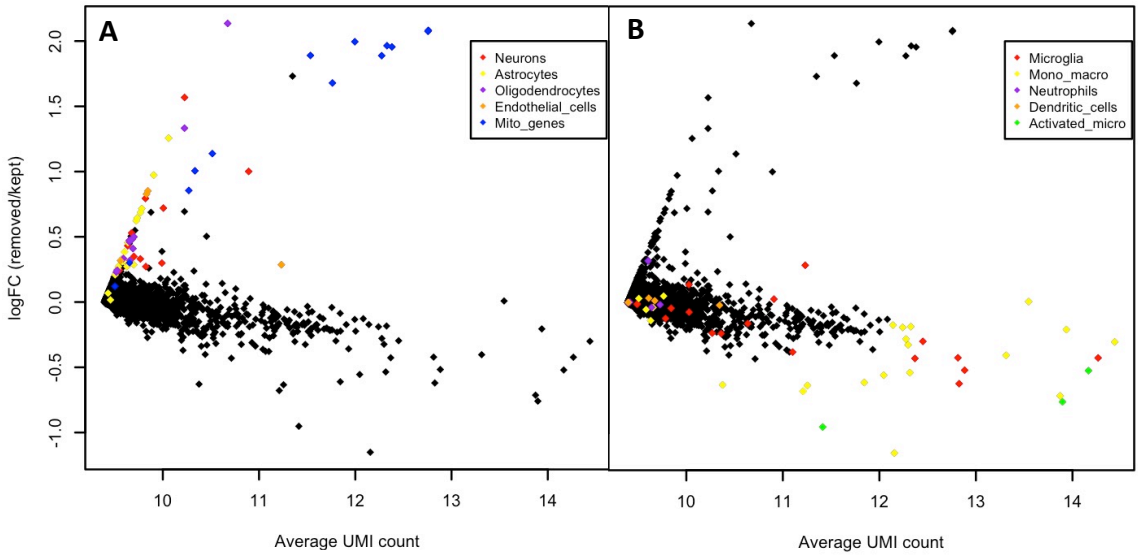


Figure 3.10 The average UMI count per gene *vs* the log fold change between the cell libraries removed or kept based on QC thresholds. Each point is a gene. Genes are coloured by the cell type in which they are highly expressed. The expression of genes from non-myeloid CNS cells is higher in the discarded cell libraries (A) while myeloid cell genes are highly expressed in the cell that remain after QC (B)

3.4.5 DOUBLET DETECTION

The 10X Genomics single cell 3' solution has a doublet rate of approximately 1% per 1000 cells. A dataset of this size likely contains a substantial volume of doublets or multiplets, which must be removed before further downstream processing to avoid confounding results. I identified doublets using the *cxds_bcds_hybrid()* function from the *scds* package (version 1.2.0) with the parameter *estNdb*l = TRUE (Bais & Kostka, 2020). Briefly, this approach predicts doublets by combining the outputs of two scoring methods: *cxds,* which evaluates the co-expression of genes across cells, looking for instances of unlikely combinations; and *bcds*, which generates artificial doubles based on the dataset, then looks for cell libraries that are similar to the artificial doublets. In total, 2,681 doublets were identified and removed from the dataset leaving 13,659 cells for downstream processing (Table 3.1). Given the number of samples (13 successful and 2 failed), this dataset is quite small. The reasons for this are discussed in *4.8*.

dpi	Sample ID	Sex	# of mice	Collected	Flow Cell	Cell Count	Median Genes/Cell	Mean Reads/Cell	Chemistry version
Ctrl	F3	M	1	Jan-19	SLX-17383	removed	249	$7.09x10^4$	2
Ctrl	H9	M	1	Mar-19	SLX-17931	removed	293	$4.28x10^4$	3
Ctrl	G11.D2	F		July-19	SLX-17850 SLX-20169	283*	1,228	$6.36x10^4$	3
Ctrl	D10	M	2	May-19	SLX-17998 SLX-20169	1,208*	1,747	$5.40x10^4$	3
1	F9	M	2	Mar-19	SLX-17931	907	1,309	$5.11x10^4$	3
1	G9.D1	M	1	Mar-19	SLX-17931	597	1,400	$6.78x10^4$	3
1	H11	M	1	Apr-19	SLX-17931	969	1,092	$5.90x10^4$	3
2	E11	M	3	July-19	SLX-17998	660	2,972.5	$3.58x10^4$	3
2	F11	F	2	July-19	SLX-17998	525	1,282	$3.16x10^4$	3
3	E3	M	1	Jan-19	SLX-17383	432	663	$1.29x10^5$	2
3	D3	M	1	Jan-19	SLX-17383	380	644	$1.75x10^5$	2
3	G9.D3	M	2	July-19	SLX-17998 SLX-20042	3,277	1,980	$5.21x10^4$	3
10	E10	M	2	May-19	SLX-17850	957	1,008	$1.14x10^5$	3
10	F10	M	1	May-19	SLX-17850	818	1,151.5	$8.44x10^4$	3
21	G11.D21	F	1	April-19	SLX-17931	841	1099	$7.95x10^4$	3

Table 3.1 The *Cx3cr1* scRNAseq samples. M=male, F=Female. # of mice indicates the number of sex, age, and condition-matched mice pooled in the scRNAseq sample. The Cell Count, and Median Genes per cell all refer to the post-QC values except for samples F3 and H9 where the C*ell Ranger* value is reported. For the samples sequenced twice to increase sequencing depth, both flow cells are reported, and the cell count, median genes/cell, and mean reads/cell are the results of the two combined sequencing runs.*after removing proliferating cells (see Normalisation).

3.4.6 NORMALISATION

Systemic variation in sequencing coverage between cell libraries can lead to confounded interpretations of scRNAseq data. To mitigate this, I normalised the dataset using a scaling normalisation, in which each cell's UMI counts are divided by a cell specific size factor. Since the *Cx3cr1* dataset comprised multiple sequencing runs, I used the *multiBatchNorm()* function from the *batchelor* package (version 1.2.4) with batches defined as the flow cell sequencing batch. I considered the samples that were sequenced twice (see Table 3.1) as their own batches for the purpose of normalisation and HVG selection. This function scales size factors relative to the average library size of each batch then uses these size factors to normalise the data, ultimately downscaling the counts in batches with higher sequencing depth to match the lowest batch. This normalisation approach eliminates an element of technical variation between the batches, making it easier to correct for batch effects downstream. The output is the log-transformed normalised UMI counts, hereon referred to as logcounts.

At this stage, I was able to investigate the normalised expression of genes of interest within the samples allowing me to perform a quick 'sanity check' to ensure that the samples exhibited expected cell markers and basic phenotypes. A more thorough investigation of the *Cx3cr1* dataset was performed after batch correction (see *3.5*) but in this 'sanity check' I noticed that a subset of the Ctrl sample, G11.D2, was proliferating (Hammond et al., 2019; Li et al., 2019; Milich et al., 2021; Tay et al., 2017). Upon closer inspection, it appeared that a few sample D10 cells were also proliferating. First, I assumed they were proliferating due to inadvertent injury during the laminectomy. However, when compared to 1 dpi and 2 dpi samples, the Ctrl samples did not show the activated microglia profile and retained the expression of homeostatic microglia genes (Figure 3.11). I also ruled out the possibility that this sample was enriched for CAMs, which are present around the Ctrl spinal cord at low levels (Goldmann et al., 2016; Jordão et al., 2019; Milich et al., 2021). It's likely that these cells were activated during the isolation process, as previously reported (Haimon et al., 2018; Hammond et al., 2019; Milich et al., 2021).

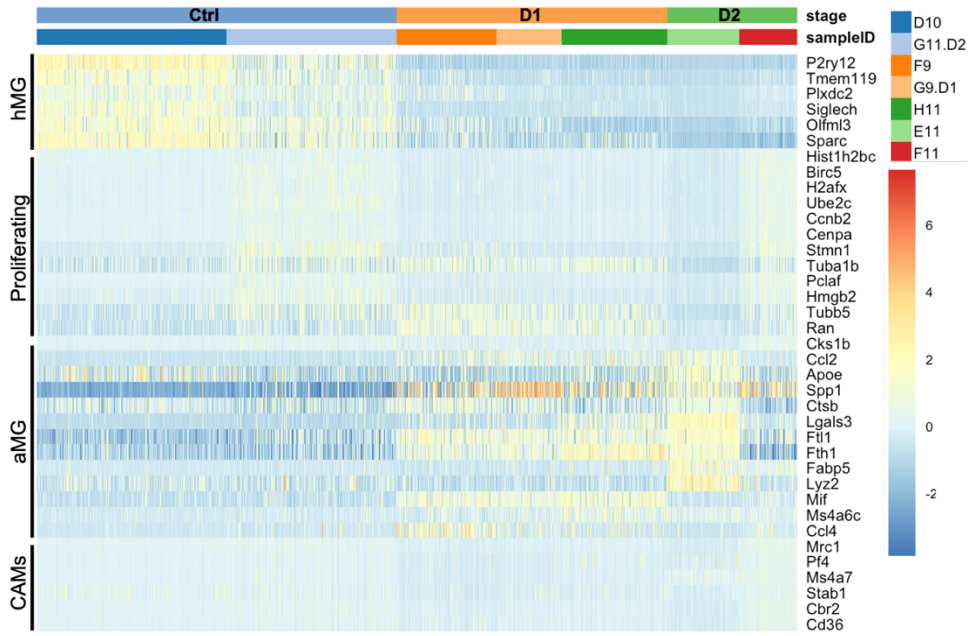


Figure 3.11 Heatmap of the Ctrl, 1 and 2 dpi samples from the *Cx3cr1* dataset. Each row is a gene, and each column is a cell. Genes are organised by the cell state they represent. hMG = homeostatic MG; aMG = activated microglia; CAMs = CNS associated macrophages. Sample G11.D2 upregulates proliferation markers, but, compared to 1 and 2 dpi samples, retains the homeostatic microglia profile, suggesting these cells are cycling but not activated.

I anticipated that this proliferating subset of sample G11.D2 would have implications for the interpretation of clustering and DEG analysis. Thus, I opted to remove these cycling cells for the downstream analyses. To this end, I clustered the cells as described in Section 3.5.1, then annotated and subset out the proliferating clusters leaving 11,852 cells in the dataset.

3.4.7 Highly Variable Gene Selection

When performing downstream analyses such as batch correction, dimensionality reduction, and clustering, considering only genes with biological, as opposed to technical, variability across cells facilitates the interpretation of the results by retaining the biological heterogeneity and, ideally, minimising the uninteresting technical noise (Amezquita et al., 2020; Lun, 2016). I identified HVGs using the *scran* package (version 1.14.6). First, I modelled the technical noise in the dataset by calling the function, *modelGeneVar,* on each batch. Batches were defined as the sequencing batch and provided through the parameter, *block*. *modelGeneVar* assumes that the variance of the majority of genes stems from the random technical noise, thus the resulting trend approximates the technical noise as a function of the genes' expression level (Figure 3.12). This trend can then be used to predict the technical and biological components of each gene's variance, where the technical component is the fitted

value and the biological component is the difference between the fitted value and the total variance (Amezquita et al., 2020). After estimating each gene's biological variance, I selected the 20% with the highest biological components as the HVGs. A total of 1,310 HVGs were selected for the downstream analyses.

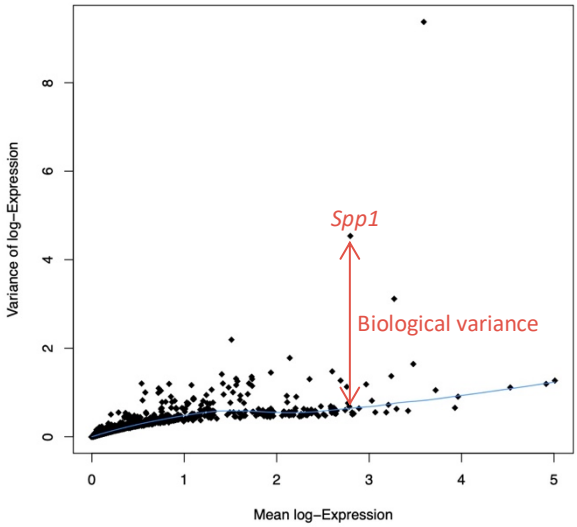


Figure 3.12 The mean log-expression *vs* variance for the 9-Apr-19 batch. Each point is a gene. The blue line is the fitted trend, which represents technical noise in the dataset. The red line with arrowheads depicts the biological variance for the gene *Spp1*.

3.4.8 Batch Correction

The *Cx3cr1* dataset consisted of five collection dates, as outlined in Table 3.2. Collecting samples in batches, while necessary for large datasets, produces technical shifts in the data that can make interpretation difficult. These shifts can be caused by a wide range of factors, such as circumstantial protocol adjustments, altered reagent quality, or different users at the Genomics Core or FACS facilities. To minimise the impact of batch effects on downstream processes, such as dimensionality reduction and clustering, we can identify and correct for these shifts. Note that for normalisation and HVG selection, I defined batches as samples that were multiplexed for sequencing as shown in *Table 3.1* and *Table 4.1*. However, for batch correction, I observed clear shifts in the UMAPs of sex- and dpi-matched samples collected on different dates but sequenced on the same flow cell (Figure 3.13A). To further validate this observation, I investigated three 1dpi samples, two of which (F9 and G9.D1) were collected on the same day, while the other (H11) was collected on different day but multiplexed with the other two. With these three samples, I performed a PCA using *scater*, then quicky clustered the cells by building a shared nearest neighbour (SNN) graph with *scran* and identified communities using the walktrap algorithm from the *igraph* package (version 1.2.6).

Then, I investigated the distribution of the three samples across the clusters and found that H11 typically clustered alone, while F9 and G9.D1 clustered together (Figure 3.13B,C). Thus, for the purposes of batch correction, I defined batch as the date of collection.

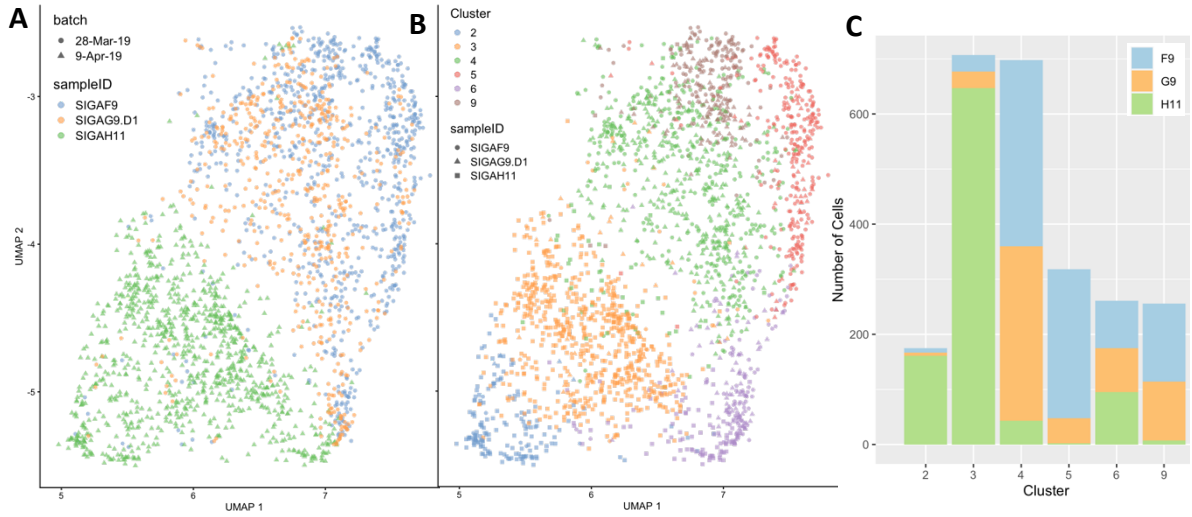


Figure 3.13 A) UMAP of three sex-matched 1 dpi samples sequenced on the same flow cell. Each cell is coloured by its sample and shaped by its collection date. We can see that sample H11 is shifted from the 1 dpi sample collected on a separate day. B) The same UMAP coloured by its cluster and shaped by its sample. C) A quantification of the cluster membership shown in (B). We can see that the two samples (G9.D1 and F9) collected on the same day also cluster together.

Since I could not assume that the composition of cell profiles was consistent across batches, the linear regression approach to batch correction was not suitable. As such, I opted for the mutual nearest neighbours (MNN) approach (Haghverdi et al., 2018) by calling *fastMNN()* from the *batchelor* package with following parameters: *batch* = the date collected, *subset.row* = HVGs, *merge.order* as shown in Table 3.2. A major assumption of the MNN approach is that there is at least partial overlap of cell populations between the two batches. Thus, an optimal experimental design would ensure that batches were arranged such that at least one condition was common between batches. It is clear from Table 3.2 that I did not have this in mind when collecting the early samples, an error I tried to mitigate with the later samples, with only partial success due to animal logistics.

Merge Order	Date Collected	dpi
1	23-May-19	Ctrl
		10
		10
2	20-Jul-19	Ctrl
		2
		2
		3
3	10-Jan-19	3
		3

4	28-Mar-19	1
		1
5	9-Apr-19	1
		21

Table 3.2 The batches comprising the *Cx3cr1* dataset, the conditions collected in each batch, and the order they were merged.

With the merge order shown in Table 3.2Table 3.2, I was able to first merge the two batches with Ctrl samples (23-May-19 and 20-Jul-19), setting a space that described the Ctrl, 2, 3 and 10 dpi populations. From here, I was forced to merge the 28-Mar-19 1 dpi samples without a guaranteed match. I reasoned that, given the experimental design, this was the best option since I expected 1 and 2 dpi to be somewhat similar and at least when merging the 9-Apr-19 1 dpi and 21 dpi samples the 1 dpi would overlap, minimising any false-correction of the sole 21 dpi sample. Satisfied with the results the *fastMNN* approach, I did not explore other batch-correction methods. However, a recent benchmarking report suggested that while *fastMNN* performed well, Harmony ranked the highest of the 14 methods under investigation (Tran et al., 2020). Harmony is a computationally inexpensive, unsupervised joint embedding method that exploits the PCA space to soft cluster analagous cells across batches whilst maximizing the variability within each cluster (Korsunsky et al., 2019). The centroid of each cluster is then used to compute a linear correction factor specific to each cluster. Each cells' factors are then averaged and applied for the correction. This process, from clustering to correction, is reiterated until Harmony converges to a stable set of clusters. Given that Harmony does not require overlapping cell populations between batches (albeit this increases the error rates) (Korsunsky et al., 2019), it may have performed even better than *fastMNN*.

After batch correction, I visualised the same 1 dpi samples on a UMAP built off the batch-corrected coordinates. I found that the three samples were now well mixed, with the exception of the infiltrating cells from sample H11, which grouped with other infiltrating cells (Figure 3.14) suggesting that the batch correction was effective. In further support of this, I investigated the variance lost within each batch during correction and found that it was ≤ 9.6%, with a mean value of 1.6%.

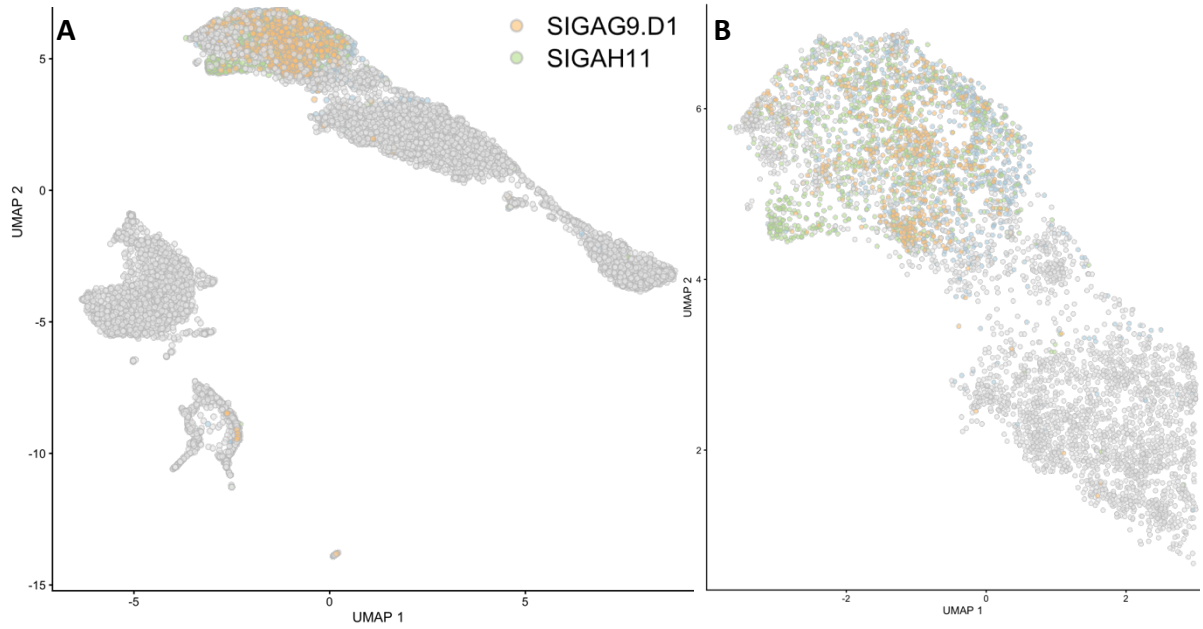


Figure 3.14 A) UMAPs of the *Cx3cr1* dataset after batch correction with the three 1 dpi samples highlighted to emphasise the correction of batch effects between collection dates compared to Figure 3.13. B) A closer look at the three 1 dpi samples from A). Each dot represents a cell Note the differences in the composition of cell types between the two batches. Sample G9.D1 contains more infiltrating myeloid cells (see Annotating Cell Types). Despite this, microglia from sample H11 were still corrected to the other microglia populations, while the infiltrating cells were not.

3.5 Initial Time-Resolved Characterisation Of Myeloid Cells Post-SCI

With the *Cx3cr1* samples pre-processed and integrated, the next step was to begin to interpret the dataset in a biologically meaningful way. I accomplished this by clustering the cells, assigning a specific myeloid cell type to each cell, and browsing through the DEGs between clusters, SCI stages, and cell types. Note that this section represents only a preliminary analysis as the *Cx3cr1* dataset was further integrated into a larger myeloid cell dataset for proper analysis.

3.5.1 Clustering

When handling data from thousands of cells, it is not feasible to investigate each cell individually. By grouping the cells together based on transcriptional similarities we can assign biological meaning to each group, making the biological interpretation of the data much more feasible. These groupings are achieved through an unsupervised learning technique called clustering. I chose to use a graph-based clustering approach, which is commonly applied to scRNAseq datasets. First, using the *scran* package, I called the *buildSNNGraph* function with the parameters *k* = 20 nearest neighbours considered, *use.dimred* = the batch corrected

dimensions, and *type* = Jaccard index. Briefly, this function builds a SNN graph where each node is a cell and the connections between nodes represent each cell's neighbour in high dimensional space. Then, using the Jaccard index, the weight of the connection between each cell is calculated, with increasingly similar cells having higher weights. After building the graph, I applied the Louvain community detection algorithm by calling the *cluster_louvain* function from the *igraph* package. This step uses modularity optimisation to identify cells that are more connected, and in theory, more similar to each other, ultimately breaking the graph into clusters (Figure 3.15A). We can evaluate the modularity between clusters as a metric of their "separateness" (Figure 3.15B), which can be useful for evaluating the biological relevance of cluster separation. Additionally, I expected some degree of modularity between cells from different dpi and especially between laminectomy-only Ctrl *vs* SCI cells. Hence, I investigated the correlation between dpi and cluster membership, and found statistically significant correlations (Figure 3.15C), suggesting a level of biological relvance for the clustering results.

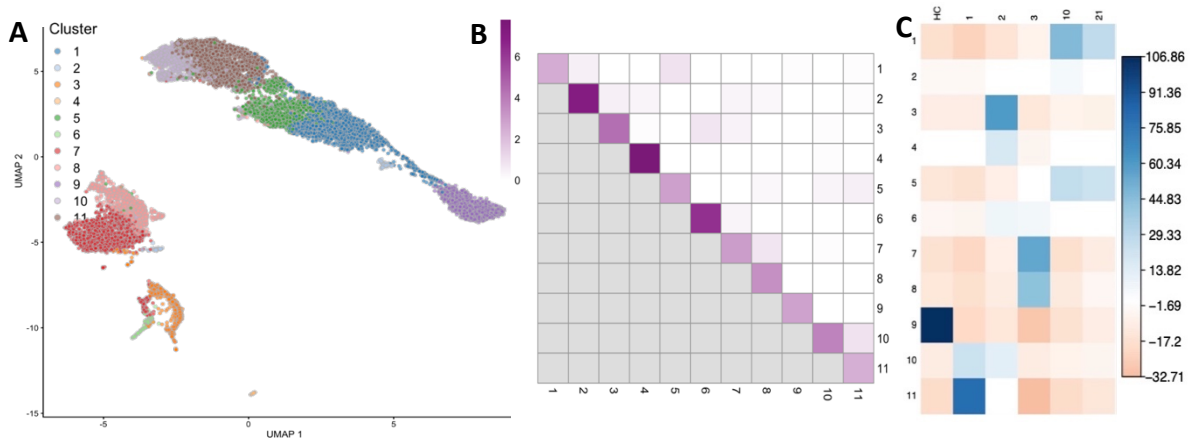


Figure 3.15 A) UMAP of the *Cx3cr1* dataset. Each dot represents a cell and is coloured by cluster membership. As we will see in the next section, these clusters can be interpreted in biologically relevant ways. For e.g., cluster 9 represents Ctrl cells and cluster 7 comprises infiltrating, but not resident myeloid cells from D3. B) A heatmap depicting the pairwise modularity score between clusters. Each row/column is a cluster. The scale bar represents the degree of modularity between clusters, compared to a null model of random connections. We can see that clusters are quite separated in general, and this separateness can have biological interpretation, such as Cluster 7/11 comprising different cell types (see Annotating Cell Types). Closely related clusters do exist, such as Clusters 10/11 or 7/8, and this too can be biologically interpreted as a similarity between cell states. C) A correlation plot of the Pearson's chi-squared residuals between cluster membership (rows) and collection stage (columns). Blue signifies a positive standardised residual, while red signifies the opposite. White represents a non-significant contribution. Certain clusters, such as cluster 9, are highly correlated with a single stage, while others, like cluster 1, are split across more than one stage. Pearson's Chi-squared test; X-squared = 61353, df = 85, p-value < 2.2e-16

3.5.2 ANNOTATING CELL TYPES

The myeloid lineage marker, *Cx3cr1*, is expressed on several distinct cell types, which I expected to find represented in the dataset. One option for annotating cell types is to simply assign a cell type label to each cluster. However, a new method, *CellAssign* (Zhang et al., 2019), provides automated cell type annotation based on cell-type marker genes without assuming that cluster membership is biologically relevant. This approach takes in a post-QC counts matrix (pre-normalisation) and a set of marker genes for the cell types of interest. It then uses a hierarchical statistical model to calculate the probability of a cell belonging to a given cell type or if the probability is below a tuneable threshold, the cell remains unassigned to minimise improper labelling.

I called the function, *cellassign,* from the *cellassign* package (version 0.99.21) using the default parameters and a custom marker gene list described in Table 3.3. The output contained maximum likelihood estimates of the cell type, which can be visualised in (Figure 3.16A). With an 80% probability threshold, 1,503 cells were still unassigned. However, these cells clearly clustered with a given cell type, justifying the use of cluster membership to manually annotate the unassigned cells (Figure 3.16B).

	Marker genes
Microglia	*Sparc, C1qa, Plxdc2, Serpine2, P2ry12, Tmem119, Siglech, Ctss, Cst3, Slc2a5, Sall1*
Macrophages	*Ms4a7, Ecm1, Arg1, Ccl7, Mgst1*
Monocytes	*Chil3, Plac8, Ccr2, Rgs1, Fn1*
Dendritic Cells	*Cd74, H2-Eb1, H2-Aa, H2-Ab1, Plac8, H2-DMb1, H2-DMa, Klrd1, Flt3, Zbtb46*
Neutrophils	*S100a8, S100a9, Lcn2, Dedd2, Retnlg, Wfdc21, Mmp9*
CAMs	*Ms4a7, Cd74, Cd163, Cbr2, Lyve1, Mrc1*
T Cells	*Cd2, Cd3d, Cd3e, Cd3g, Ptprc*

Table 3.3 The manually curated marker genes used for *CellAssign*. The markers draw heavily from previous scRNAseq myeloid cell studies (Jordão et al., 2019; Milich et al., 2021; Zeisel et al., 2018) and PanglaoDB.

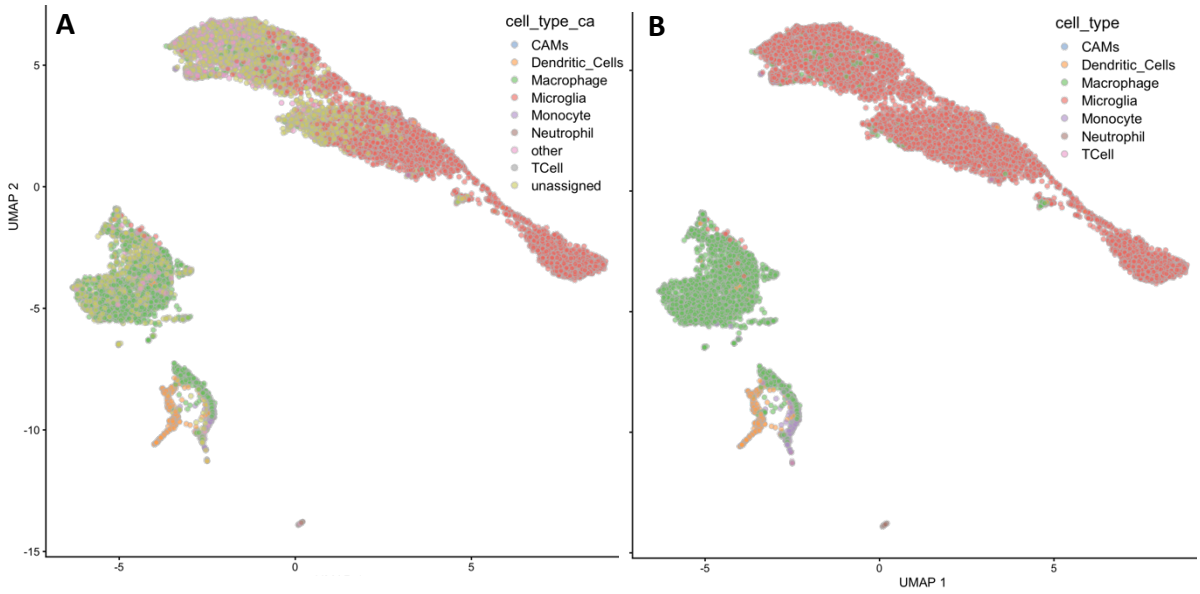


Figure 3.16 UMAPs of the *Cx3cr1* dataset coloured by *CellAssign* estimated cell type (A). In B) the "other" and "unassigned" cells have been manually annotated based on the majority *CellAssign* cell type of their cluster. We can see that specific cell types tend to group together on the UMAP. However, there is some overlap between microglia and macrophages warranting further inspection.

At this stage, I noted that several myeloid cell types were present in the dataset, and as expected, this changed over time (Figure 3.17A). The trends for infiltrating macrophages and neutrophils are in line with previous studies (Donnelly & Popovich, 2008; Stirling & Yong, 2008). Dendritic cells were previously reported to infiltrate the spinal cord only at 56 dpi (Sroga et al., 2003) but more recently have been identified in the acute phase of T8 contusion SCI in WT mice (Milich et al., 2021), in line with my observations. Of note, in several cases, the distinction between resident microglia and infiltrating myeloid cells was difficult to discern based on transcriptional profiles alone, resulting in both cell types being present in the same cluster (Figure 3.17B, C), assuming the cell type annotation was correct.

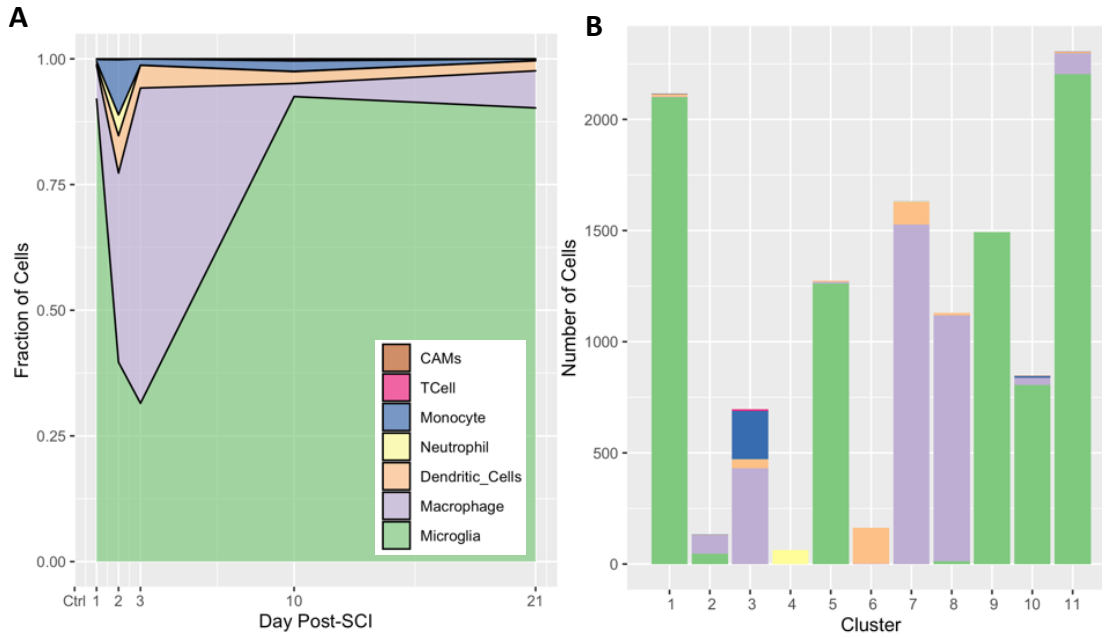


Figure 3.17 A) Area composition plot of the fraction of cells captured at each stage of SCI. The dynamics are in line with previous reports. B) Bar chart depicting the number of cells per cluster, coloured by their annotated cell type as in (A). In some cases, clusters are cell type specific, but not always. However, this observation relies on accurate cell type annotation.

3.6 DISCUSSION

In this chapter I have described the methodology used to isolate and sequence *Cx3cr1*$^{+}$ myeloid cells, the workflow I built to compile this myeloid cell atlas, drawing heavily on the literature for the design and justification of each step, and finally, I have performed a preliminary analysis of the myeloid cell atlas.

The use of the *Cx3cr1* mouse allowed for the efficient isolation of myeloid cells from the spinal cord. This targeted approach ultimately allowed us to investigate myeloid cell states with high resolution, revealing heterogeneity that might otherwise have been overshadowed by other CNS cell types. However, by enriching for certain cell populations before scRNAseq, we lose the ability to investigate these cells in the context of the system through, for example, the inference of intercellular networks via ligand-receptor analysis (Efremova et al., 2020). An optimal approach, in the absence of financial constraints, would be to divide the sample, enrich half for the population of interest and sequence the other half in its entirety.

After applying the QC thresholds, I examined the average counts per genes for the kept *vs* discarded cells (Figure 3.10) and found that the discarded cells upregulated genes of non-myeloid CNS cells. Since we isolated *Cx3cr1*$^{+}$ cells using FACS, it is unlikely that large quantities

of *Cx3cr1*$^{-}$ cells were present in the sample. Rather, given that the samples were extracted from tissue via mechanical homogenisation, it is plausible that myeloid cells remained largely intact, while more fragile CNS cells such as neurons and astrocytes were damaged and released their mRNA into the cell suspension. These large quantities of ambient mRNA would have then been barcoded in empty droplets, ultimately generating what appeared to be small cell libraries with low complexity but likely artefacts. Supporting this theory, and the efficacy of the QC thresholds, the genes upregulated in the retained cell libraries were myeloid cell genes.

Aside from mitochondrial genes, some of the genes with the highest FC in the discarded libraries were neuronal genes. Notably, there have also been reports that the *Cx3cr1* mouse line I employed has leakage into ~5% of neurons (Zhang et al., 2018), although it's unclear if this finding was biased by the author's choice of reporter mouse (Zhao et al., 2019). Additionally, there are reports of activated-neuronal *Cx3cr1* expression (Dworzak et al., 2015; Wang et al., 2018). Consequently, if a leakage was present or *Cx3cr1* was expressed in neurons post-SCI, YFP^{+} neurons could have been barcoded and sequenced. In that case, as previously mentioned, neurons are more fragile than myeloid cells and would very likely be damaged during the cell isolation protocol, and thus wouldn't pass the QC stage.

The *Cx3cr1* discussions above exemplify how experimental design has critical implications for the *in silico* processing of scRNAseq data. Another instance, crucial for any experiment requiring multiple collection dates, is the careful planning of samples across batches. Ideally, biological replicates should be spread across batches to improve the accuracy of detecting and correcting for batch effects. As mentioned in Section 3.4.8 on Batch Correction, this was not taken into consideration for the first two batches in this dataset and, while every attempt was made to compensate for this with later batches, the MNN merging of batch 28-Mar-19 lacked a guaranteed overlap between batches. Thus, it is possible that 1 dpi cells were erroneously mapped to cells from other stages. However, Figure 3.17C suggests not, as only cluster 10 showed significant correlation to both 1 and 2 dpi. Regardless, thoughtful experimental design before collecting samples for scRNAseq is essential.

Knowing that myeloid cells are particularly sensitive to perturbations, including during scRNAseq preparation (Haimon et al., 2018), I carefully designed the myeloid cell isolation protocol and was able to demonstrate that this protocol induced fewer dissociation- and FACS-associated genes than previous microglia studies. Yet, even with these efforts, there were still signs of isolation-induced transcription. Since completing this study, a new protocol has been released (Marsh et al., 2020), which employs transcription and translation inhibitors to prevent *ex vivo*-induced transcription even when enzymatic dissociation is used. Unfortunately, the scRNAseq data from this study has not yet been released, so I was unable to compare their results with my own data. However, regarding this study, employing such a protocol might have had benefits beyond reducing *ex vivo*-induced transcription. Primarily, it would have likely increased the quality of the isolated myeloid cells and reduced the volume of ambient mRNA, ultimately resulting in a cleaner, larger myeloid cell atlas.

The myeloid cell atlas presented in this chapter comprised most of the expected cell types, with fluctuating prevalence as described by previous studies (Donnelly & Popovich, 2008). Notably absent from the dataset were CAMs (*CellAssign* annotated 4 CAMs). These cells have been described in the spinal cord (Bellver-Landete et al., 2019; Jordão et al., 2019; Plemel et al., 2020), but their absence in this dataset can likely be attributed to the use of the spinal cord hydraulic extrusion method, which removed the meninges. A recent scRNAseq study of all CNS infiltrating and resident cells post-SCI isolated the spinal cord through dissection and identified small quantities of CAMs in their dataset at 1, 3, and 7 dpi (Milich et al., 2021). However, they did not investigate potential roles for these cells in the SCI pathology. Another recent study using bone marrow chimeras found that CAMs infiltrated the injured spinal cord and contributed to the microglia scar at 14 dpi (Bellver-Landete et al., 2019). Whether chimeras were suitable for this type of study is unclear, since irradiation is well known to compromise BBB and BSCB integrity and alter the function of resident myeloid cells (Bruttger et al., 2015; Mildner et al., 2011). If CNS infiltrated CAMs are present in the spinal cord after SCI, a possible justification for their absence in this myeloid cell atlas is that given their low quantities, there was a low chance of capturing *Cx3cr1*$^{+}$ CAMs compared to the dominant macrophage or microglia populations. Arguably, this effect was observed with the neutrophil population, where only 65 cells (0.5% of the dataset) were present. In the context of EAE, CAMs have been argued to regulate the infiltration of immune cells into the CNS (Engelhardt

et al., 2001; Kivisäkk et al., 2009; Reboldi et al., 2009), and although this finding was recently contradicted (Jordão et al., 2019), it remains possible that CAMs could contribute to the perpetual inflammation observed in chronic SCI. Ultimately, the role, or lack thereof, for CAMs after SCI has yet to be fully elucidated.

Aside from CAMs, this myeloid cell atlas captured the anticipated resident and infiltrating myeloid cell populations. The most prevalent infiltrating myeloid cells were macrophages, which in some cases clustered with stark modularity (e.g., cluster 8) but also appeared to cluster with certain microglia populations (e.g., clusters 2/11). Recent experiments have demonstrated that CNS infiltrating macrophages take on the morphology and transcriptional profiles of microglia (Bennett et al., 2018; Greenhalgh & David, 2014), inciting the question – was I capable of accurately annotating infiltrating macrophages in this dataset? To fully address this, I had to re-evaluate my experimental design.

Chapter 4 Disentangling Resident *vs* Infiltrating Myeloid Cells by Employing a Fate mapping Mouse Model of SCI

"La vie n'est facile pour aucun de nous. Mais quoi, il faut avoir de la persévérance, et surtout de la confiance en soi. Il faut croire que l'on est doué pour quelque chose, et que, cette chose, il faut l'atteindre coûte que coûte."

— Marie Curie

Disclosure

All methods were planned, trialled, and performed by me, unless otherwise stated. Yutong Chen, an undergraduate medical student from the University of Cambridge, assisted me with some of the *ex vivo* procedures and Veronica Testa, a visiting master's student from the University of Milan Bicocca, helped me with some of the *ex vivo* procedures and performed approximately half of the tissue preparation for immunofluorescence (IF) and smFISH.

Certain aspects of this chapter are expanded from the bioRxiv pre-print: Hamel, R. *et al.* Time-resolved single-cell RNAseq profiling identifies a novel Fabp5-expressing subpopulation of inflammatory myeloid cells in chronic spinal cord injury. *bioRxiv*, 2020.2010.2021.346635, doi:10.1101/2020.10.21.346635 (2020). I wrote the manuscript for this pre-print in its entirety. Luca Peruzzotti-Jametti, Stefano Pluchino, and John Marioni edited the manuscript.

4.1 INTRODUCTION

As discussed in, *1.1.3 A Closer Look at the Immune Response to SCI*, myeloid cells represent a diverse set of immune cells that respond to tissue damage and pathogens (Ransohoff & Cardona, 2010). Within the myeloid cell family there are distinct ontologies. Microglia and most CAMs arise from extra-embryonic yolk-sac during development and migrate to the neonatal brain before the formation of the BBB. They maintain their population throughout adulthood via local proliferation in the CNS parenchyma (Ginhoux et al., 2010; Goldmann et al., 2016). Microglia and CAMs are generally long-lived, with a lifespan of several years in humans (Reu et al., 2017) and over a year in mice (Tay et al., 2017). Conversely, infiltrating neutrophils and monocyte-derived myeloid cells, such as macrophages and dendritic cells, arise from HSC in the bone marrow and spleen (Milich et al., 2019). Under healthy conditions, these cells are very short lived (1-3 days) (Kezic & McMenamin, 2008). Under pathological conditions, such as SCI, these peripheral myeloid cells are honed to the injury site. In the case of monocyte-derived macrophages, after crossing the BBB or BSCB, they can take on the morphology of microglia, and upregulate canonical microglia markers (Bennett et al., 2018). Given that activated microglia downregulate their homeostatic signatures, these two cell populations can become practically indistinguishable (Butovsky et al., 2014; Dubbelaar et al., 2018).

In this chapter, I employ a fate mapping transgenic mouse line to unequivocally isolate *Cx3cr1*$^{+}$ microglia and infiltrating myeloid cells from mouse models of contusion SCI. I collected samples at identical timepoints and used the same protocols and workflows that I established for the previous *Cx3cr1* dataset (*Chapter 3*). This allowed me to ultimately integrate these datasets into a time-resolved myeloid cell atlas with less ambiguous annotations of cell ontogeny, which I then used for further analysis. I also generated a publicly available, interactive web app, where other researchers can explore and utilise this data. The code I wrote for handling the data in this chapter is publicly available on my GitHub (https://github.com/regan-hamel/PhD_Dissertation).

4.2 CHARACTERISING THE FATE MAPPING MOUSE MODEL

To separate infiltrating from resident myeloid cells, I adopted a transgenic fate mapping mouse line dubbed *CreMato.* To generate this line, with the help of Luca Peruzzotti-Jametti, a post-doctoral researcher in the Pluchino lab, and Daniel Trajkovski, a lab technician, I crossed homozygous *Cx3cr1* mice used in *Chapter 3* with homozygous B6.Cg-Gt(ROSA)26Sor$^{tm9(CAG-tdTomato)Hze/J}$ mice (Madisen et al., 2010) carrying a STOP cassette flanked by two loxP sites preventing the expression of the tdTomato transgene until tamoxifen administration. The latter strain is referred to hereon as tdTomato or tdTomato^{flox}. The resulting *CreMato* mice (Jordão et al., 2019) express YFP in all *Cx3cr1*$^{+}$ cells and, when treated with 5 consecutive days of tamoxifen (intraperitoneal injections of 125 mg/kg of body weight dissolved in corn oil; Sigma-Aldrich T5648-5G), *Cx3cr1*$^{+}$ cells become double positive for YFP and tdTomato.

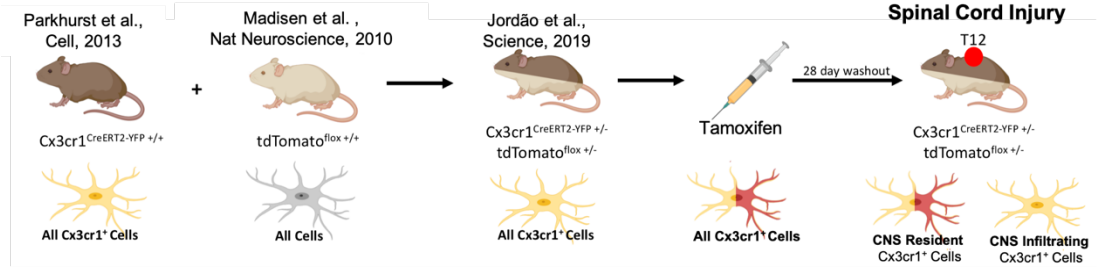


Figure 4.1 A schematic depicting the generation of the *CreMato* mouse, in which tamoxifen treatment and a 28-day washout period leaves resident myeloid cells double positive for YFP and tdTomato, and infiltrating myeloid cells are single positive for YFP.

As previously described, resident myeloid cells are self-renewing, while peripheral myeloid cells are continuously generated *de novo* from HSC. Thus, the latter will only transiently express tdTomato. After a 28-day washout period, resident myeloid cells, such as microglia, remain double positive for tdTomato and YFP, while infiltrating myeloid cells are single positive for YFP, allowing these cells to be easily separated by FACS and other fluorophore-based techniques including IF microscopy (Figure 4.1).

Given that the *CreMato* mouse line was a pivotal tool employed in *Chapters 4* and *5* to distinguish between infiltrating *vs* resident myeloid cells, with the help of several Pluchino lab members and collaborators, I confirmed that the fluorophore expression labelled the expected cell-types.

4.2.1 TISSUE PROCESSING

Mouse models of T12 contusion SCI were generated as described in *0*. At 1, 2, 3, 10, or 21 dpi, I deeply anaesthetised the mouse with an intraperitoneal injection of 100 µl of pentobarbital sodium before performing transcardial perfusion with carbogen-bubbled ice-cold artificial cerebral spinal fluid (aCSF) for 7 minutes or until the liver was clear of blood. Then, I perfused the mice with ice-cold 4% paraformaldehyde (PFA; Sigma Aldrich 441244) in PBS with a pH 7.4 for 7 minutes or until the tissue was very stiff. Afterwards, I dissected out the spinal column and post-fixed it overnight in the same PFA solution.

The following steps were performed by either Veronica Testa, a visiting master's student from the University of Milan Bicocca, or me, in line with the lab's published protocols (Braga et al., 2020). We dissected the spinal cord from the column, washed it in PBS, and then cryoprotected it in 30% sucrose (Fischer Scientific 10634932) until the tissue no longer floated in the solution (24-48hrs). Then we cut the spinal cord to a length of 15 mm centred on the laminectomy site and embedded it in OCT embedding compound (CellPath KMA-0100-00A). We snap froze the spinal cord by lowering the tissue into an OCT-filled mould (Millipore Sigma E6032-1CS) and then placed the mould into an isopentane bath filled with dry ice. We stored the OCT-tissue blocks at -80˚C until use.

To section the tissue, Veronica or I used a cryostat (Leica CM1850) to cut the spinal cord into coronal sections of 20 µm thickness. The sections were collected onto Superfrost Plus (ThermoScientific 10149870) or Superfrost Plus Gold slides (ThermoScientific K58000AMNZ72) in a series of 20 slides. Each slide represented the entire 15 mm spinal cord section, with the tissue sampled every 400 µm. We stored the slides at -80˚C until use.

4.2.2 SINGLE-MOLECULE FLUORESCENT *IN SITU* HYBRIDIZATION

Katherine Ridley, a lab technician in the Rowitch lab, Stefano Pluchino, and I collectively designed the smFISH experiments. Katherine performed these experiments in their entirety and her methodology is described in the pre-print (Hamel et al., 2020).

We used smFISH to confirm that non-myeloid CNS cells did not express YFP or tdTomato. We investigated Ctrl, 1 dpi, and 21 dpi tissue at the lesion epicentre and above and below the

injury site (Figure 4.2). Katherine did not observe the expression of markers from non-myeloid CNS cell types, indicating that the *CreMato* mouse specifically labelled our target population. Notably, this smFISH data was not quantified.

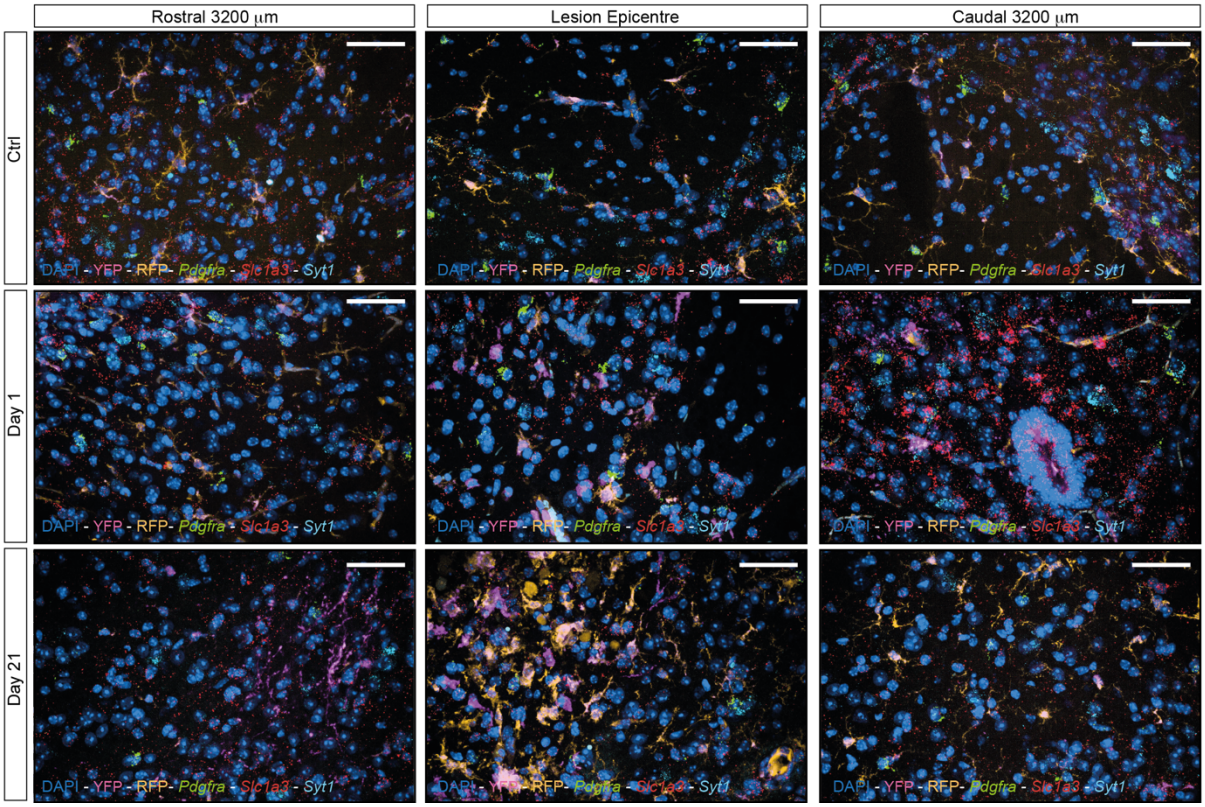


Figure 4.2 smFISH Expression of cell-type markers *Pdgfra* (oligodendrocyte progenitor cells), *Slc1a3* (astrocytes), and *Syt1* (neurons) as well as IF labelling YFP and tdTomato (RFP). We did not observe co-localisation between RFP and/or YFP in any of the investigated cell types. Nuclei are stained with DAPI. Scale bars: 60 µm.

4.2.3 Immunofluorescence & Quantification of Infiltrating *vs* Resident Myeloid Cells

To confirm the myeloid-lineage specificity of the *CreMato* mouse line, Luca performed the IF staining and confocal IF microscopy for Figure 4.3, as described (Hamel et al., 2020). As anticipated, Ctrl samples comprised $YFP^+/tdTomato^+$ microglia only but upon SCI, CNS infiltrating $YFP^+/tdTomato^-$ myeloid cells, namely monocytes/macrophages, neutrophils, and myeloid dendritic cells, increased considerably. By 21 dpi, infiltrating myeloid cells remained present with the exception of neutrophils, as previously described (Donnelly & Popovich, 2008).

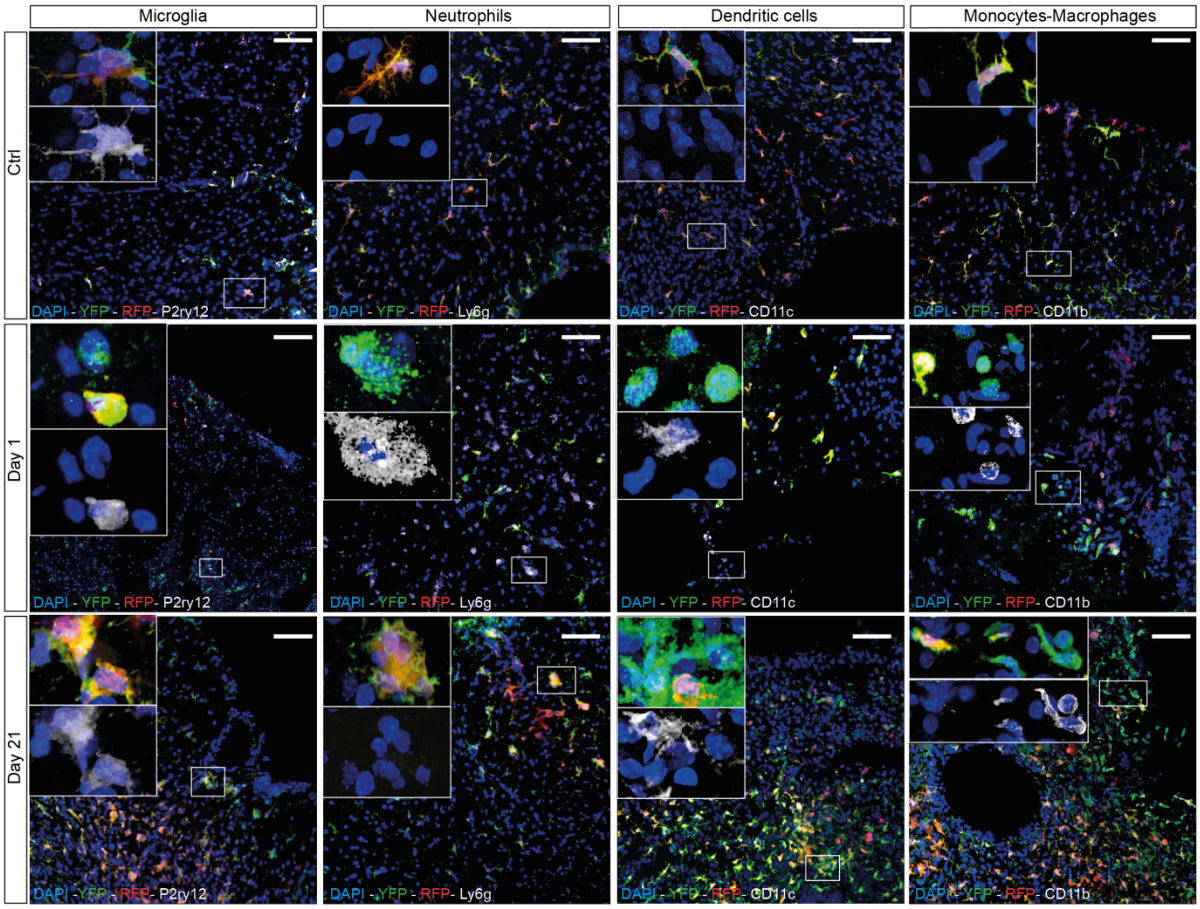


Figure 4.3 Confocal IF microscopy at the lesion epicentre demonstrating the colocalization of the microglia marker, P2ry12, with YFP and tdTomato, and the colocalization of infiltrating myeloid cell markers Ly6g (neutrophils), CD11c (dendritic cells) and CD11b (monocytes/macrophages) with YFP only. In line with the literature and scRNAseq data, we found that Ctrl tissue comprised microglia only. Upon SCI, CNS infiltrating myeloid cells were present at the lesion epicentre at 1 and 21 dpi, except neutrophils which were absent at 21 dpi. Nuclei were stained with DAPI. Scale bars: 60 µm.

A major caveat of scRNAseq is the loss of spatial context. In this study, I isolated myeloid cells from a 5 mm section of tissue centred on the lesion. Thus, any information regarding the number of resident or infiltrating myeloid cells at the lesion epicentre *vs* the spared tissue above or below the lesion was lost. To recover a small proportion of that context, Veronica, and I, with advice from Luca and Stefano, designed an IF-based experiment to quantify the spatial distribution of infiltrating *vs* resident myeloid cells over time post-SCI.

The majority of the IF staining and microscopy, and the entirety of the quantification, was performed by Veronica. I helped her with about a quarter of the staining and microscopy, and I worked closely with her to construct the automated and manual quantification pipelines. Luca Peruzzotti-Jametti provided input on the staining protocol. The IF, microscopy, and quantification protocols are described in the pre-print (Hamel et al., 2020). Briefly, after

staining for YFP and counter staining with DAPI, we imaged the tissue sections at 20X using an epifluorescence microscope (tdTomato was robust enough to be detected without staining). Then, Veronica identified two regions of interest per section, one on either side of the necrotic lesion core, or similar anatomical region for the spared tissue sections (Figure 4.4). Finally, she used an automated pipeline to quantify $DAPI^+$ cells in the regions of interest, and in the same regions she manually counted YFP+/$tdTomato^+$ resident and YFP^+/$tdTomato^-$ infiltrating myeloid cells.

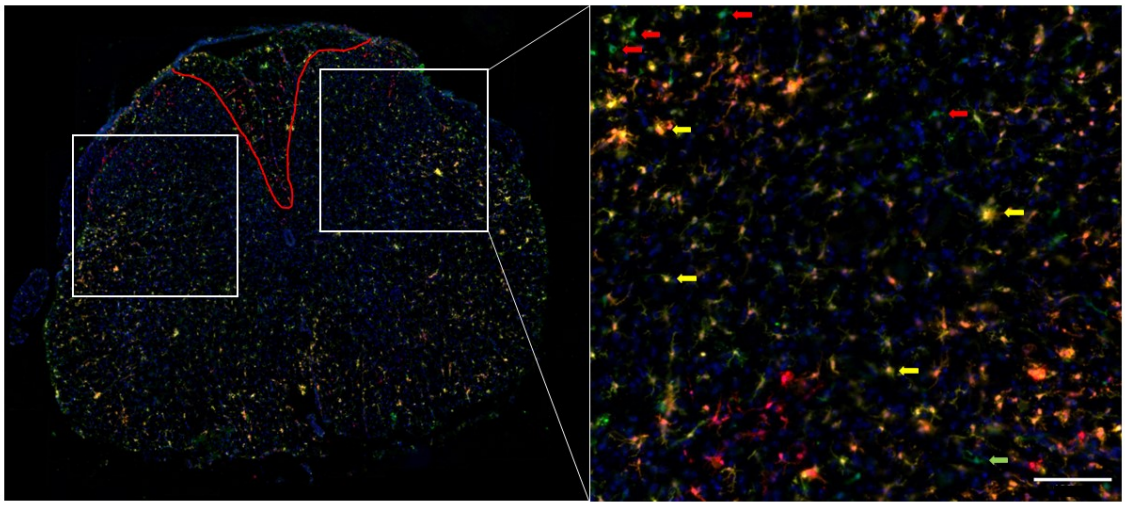


Figure 4.4 Representative images from Veronica of the IF used for the quantification of resident *vs* infiltrating myeloid cells. On the left, the coronal spinal cord section stained for YFP and DAPI (the tdTomato signal was strong enough to be visualised without staining). The red line outlines the necrotic lesion core, and the white boxes represent the regions of interest Veronica selected for her quantifications. On the right, a representative region of interest. The red and yellow arrows are infiltrating and resident myeloid cells, respectively. Scale bar: 100 µm

Using this method, we found that resident myeloid cells increased over time and relative to the proximity to the lesion epicentre (Figure 4.5A), in line with previous findings (Bellver-Landete et al., 2019). We observed a similar trend in infiltrating myeloid cells but with a stark increase in infiltrating myeloid cells at the lesion epicentre at 21 dpi compared to the acute and subacute time points (Figure 4.5B).

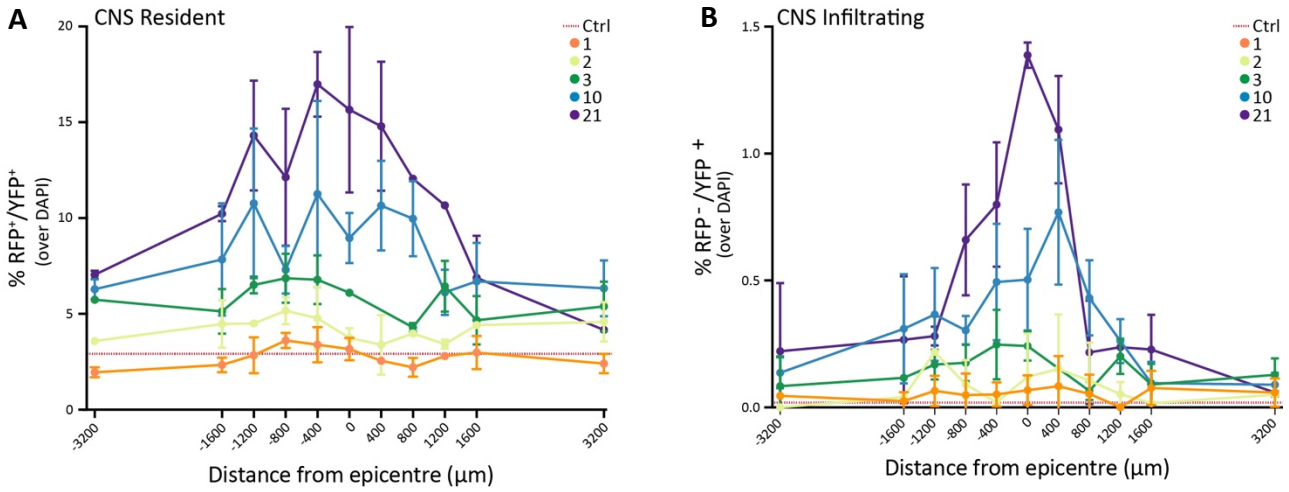
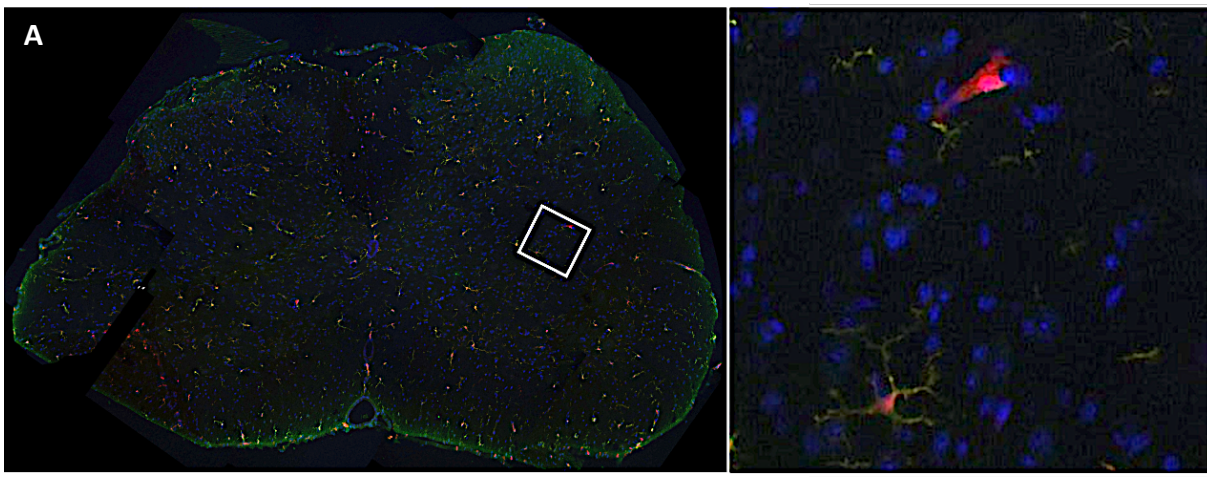


Figure 4.5 Quantification of A) YFP⁺/tdTomato⁺ CNS resident myeloid cells or B) YFP⁺/tdTomato⁻ CNS infiltrating myeloid cells as a proportion of DAPI⁺ cells across set distances rostral (+) or caudal (-) to the lesion epicentre. Each graph is coloured by the dpi. Data are mean % (± SD) and n = 3 per dpi.

In contrast to the smFISH images, when imaging this *CreMato* tissue, Veronica and I observed many instances of YFP⁻/tdTomato⁺ cells with distinctly non-myeloid morphology (Figure 4.6A). This suggested that tdTomato can be expressed in a cre recombinase-independent manner, which was recently described (Jordão et al., 2019; Zhao et al., 2019). I also noticed that at later timepoints, such as 21 dpi, some resident myeloid cells, particularly at the lesion epicentre, highly expressed tdTomato but barely expressed YFP (Figure 4.6B). Given that activated microglia proliferate through clonal expansion (Tay et al., 2017) and downregulate *Cx3cr1* (Dubbelaar et al., 2018), these YFP^{lo}/tdTomato⁺ cells could represent activated resident myeloid cells derived from clonal expansion after SCI.



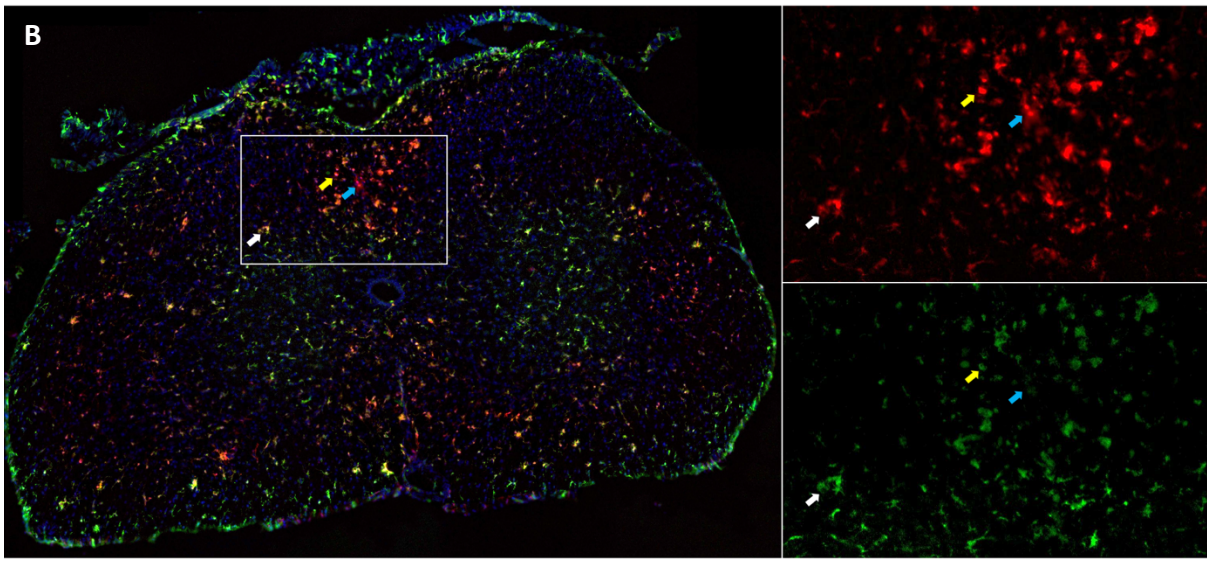


Figure 4.6 A) Ctrl *CreMato* spinal cord section (left) with a magnification (right) depicting a YFP$^+$/tdTomato$^+$ ramified microglia and a much larger YFP$^-$/tdTomato$^+$ (presumably neuronal) cell, confirming that the tdTomato bleeds through to other cells without an active cre-recombinase. B) 21 dpi *CreMato* spinal cord section (left) with magnification of tdTomato and YFP channels, respectively (right). Arrows indicate double positive resident myeloid cells of interest. White arrows depict a cell with high expression of both YFP and tdTomato. Yellow arrows depict a cell with high tdTomato and low YFP expression. Blue arrows depict a cell with high tdTomato and nearly indistinguishable YFP expression.

4.3 Infiltrating *vs* Resident Myeloid Cell Isolation

In this chapter, I prepared the tissue samples for scRNAseq using an identical protocol to *3.2 Myeloid Cell Isolation*, except for the FACS. For this step, I once again worked with the technicians at the Cambridge Phenotyping Hub to isolate live myeloid cells (DAPI$^-$ YFP$^+$) using a BD FACS Aria III or Influx cell sorter. As before, we set the sorter to 3-way purity and 20 psi with a 100 µm nozzle. For the *CreMato* samples, the panel consisted of only YFP, tdTomato, and DAPI, so fluorochrome compensation was not required. We set the gates based on an unstained WT sample and a *CreMato* sample without tamoxifen treatment to evaluate the level of cre recombinase-independent tdTomato leakage (Figure 4.7A). In line with previous reports, we observed leakage of tdTomato into YFP$^+$ and YFP$^-$ cells, in the absence of tamoxifen treatment (Jordão et al., 2019; Zhao et al., 2019). As in our IF experiments, we observed YFP$^-$/tdTomato$^+$ cells. As with the *Cx3cr1* samples, to increase cell yield but minimise the isolation of non-myeloid cells or the mis-categorisation of infiltrating *vs* resident myeloid cells, we set the cell and singlet gates generously, but were strict on the live, YFP$^+$/tdTomato$^+$ and YFP$^+$/tdTomato$^-$ gates (Figure 4.7B).

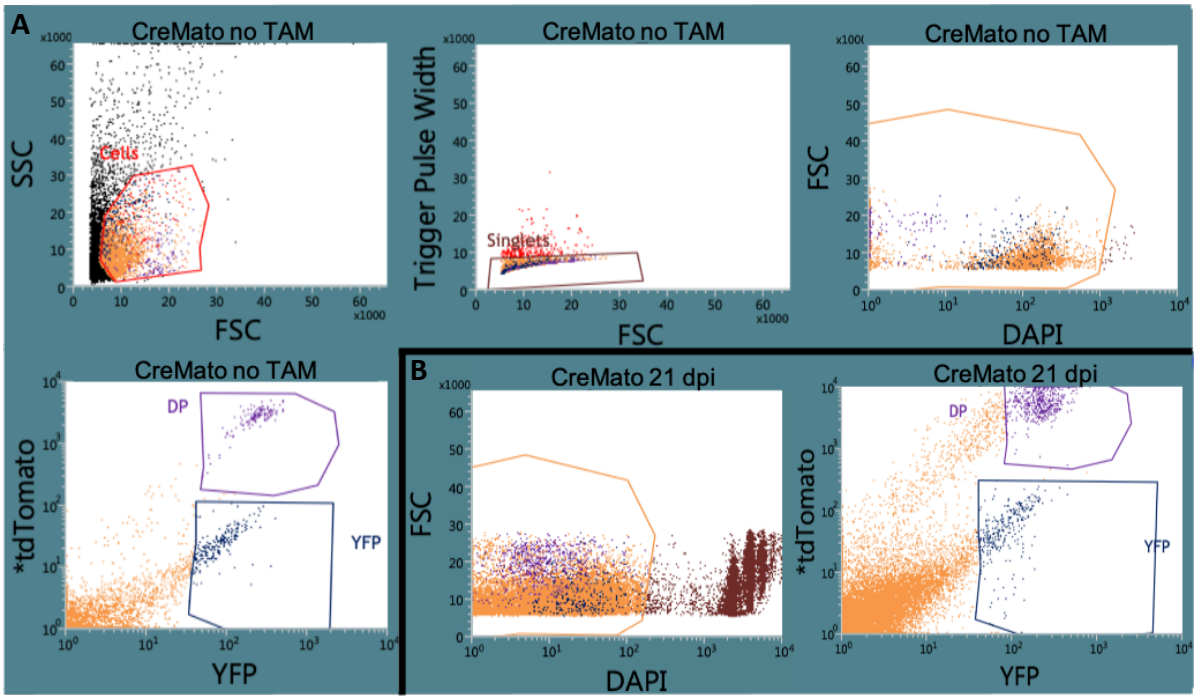


Figure 4.7 A) representative FACS gating strategy based on an unstained *CreMato* sample with no tamoxifen treatment. Note that even without tamoxifen treatment, some cells are double positive or YFP$^-$/tdTomato$^+$. B) I applied the gates from (A) to a 21 dpi *CreMato* sample stained for DAPI. SSC = side scatter; FSC = forward scatter.

4.4 Single-cell RNA Sequencing

The staff at the CRUK Cambridge Institute Genomics Core sequenced the cells with single-cell acuity using a microdroplet based-platform, 10X Genomics Chromium Single Cell 3' Solution followed by Illumina sequencing. All protocols were performed as described in *3.3 Single-Cell RNA Sequencing* and all samples were processed with version 3 chemistry.

4.5 Compiling the Time-Resolved Infiltrating *vs* Resident Myeloid Cell Atlas

4.5.1 Pre-processing

For the *CreMato* dataset, I performed each of the pre-processing steps as described in *3.4 Compiling the Time-Resolved Myeloid Cell Atlas*. Briefly, I used *CellRanger* to align the reads to the edited mm10 mouse genome and then performed all remaining scRNAseq data analysis in R version 3.6.3. All the packages I used are publicly available and my complete R analysis workflow for this chapter is available on GitHub: https://github.com/regan-hamel/PhD_Dissertation. From the *CellRanger* unfiltered feature-barcode matrix, I corrected for barcode swapping, performed cell calling, and then used this filtered feature-barcode matrix for low-level analysis and further downstream processing.

4.5.2 Quality Control

As discussed in *Chapter 3*, low quality cells must be removed from the dataset before further analysis and biological interpretation. To this end, I called the *perCellQCMetrics* function from the *scater* package to calculate the QC metrics for each cell. Like the *Cx3cr1* dataset, in the *CreMato* dataset, the distributions of library sizes within each sample were similar to the complexity, and in many cases cells with smaller library sizes and lower complexity had higher proportions of mitochondrial genes (Figure 4.8).

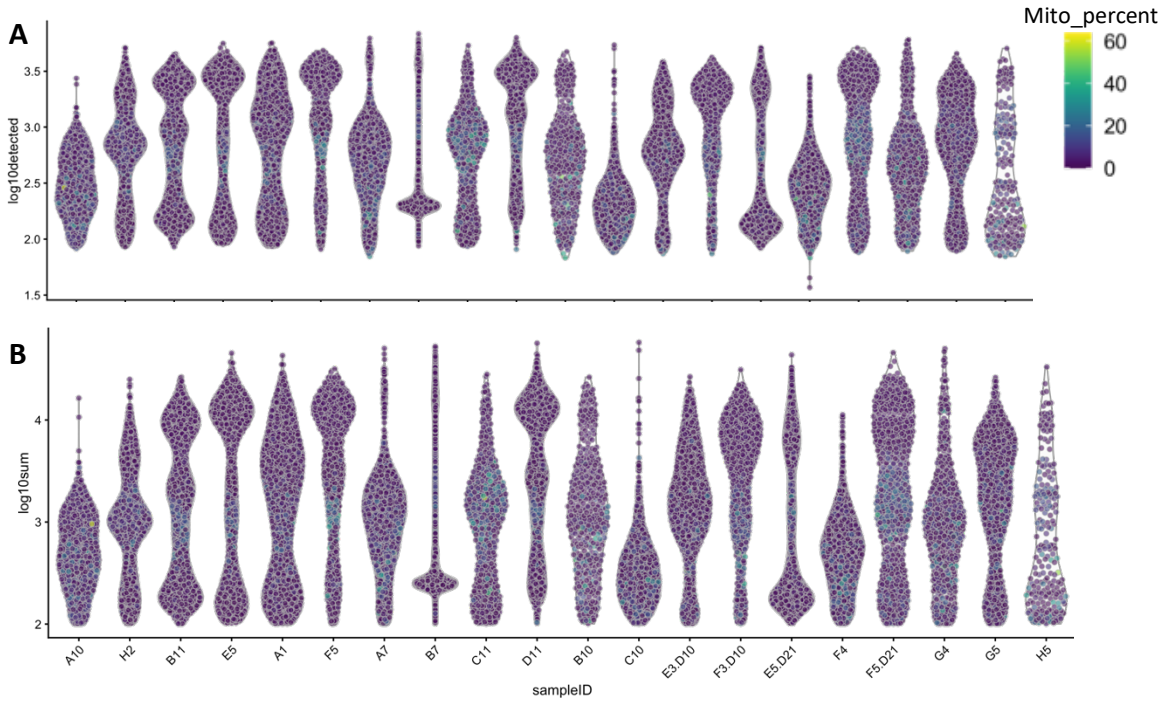


Figure 4.8 Violin plots depicting the log10 number of detected genes per cell (A) or UMI counts per cell (B) for each *CreMato* sample. Each cell is coloured by its proportion of mitochondrial transcripts. As in the *Cx3cr1* dataset, cells with fewer detected genes and fewer UMI counts often expressed higher proportions of mitochondrial transcripts, suggesting these cells are of low quality.

As with the *Cx3cr1* dataset, I checked whether the variation of the QC metrics observed across the samples might correspond to the batch (Figure 4.9A). For the *CreMato* dataset, this was generally true for samples with the same fate mapping label (Figure 4.9B). On average, infiltrating cell libraries had higher UMI counts (3,359.0 *vs* 2,235.0) and more complexity (938.4 *vs* 778.5) than resident cells.

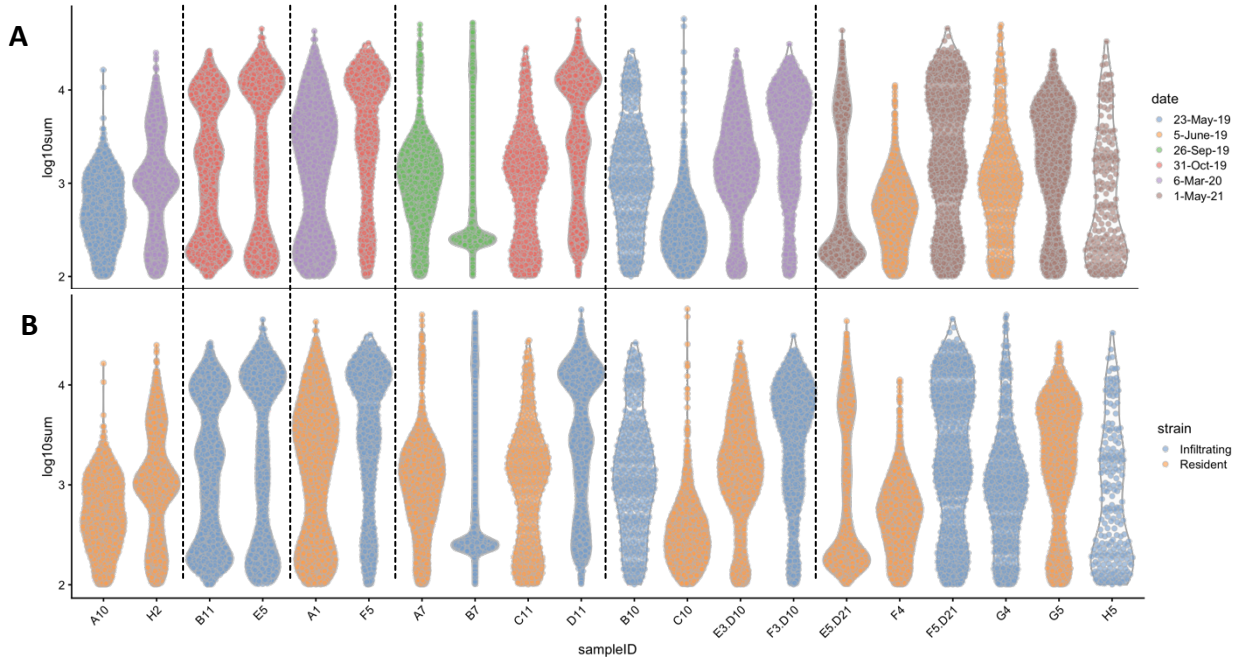


Figure 4.9 Violin plots depicting the log10 UMI counts per cell for each *CreMato* sample. Each cell is coloured by the batch (A) or the fate mapping label (B). Samples from the same condition that were collected on the same day appear to have similar distributions of cell library sizes. Dotted lines separate the samples by dpi in the following order: Ctrl, 1, 2, 3, 10, 21 dpi

Next, I applied the same QC thresholds to the *CreMato* dataset as in *3.4.4 Quality Control*. Briefly, I used fixed thresholds of 5% mitochondrial transcripts (Figure 4.10), 1,000 UMIs and 600 genes per cell. After applying these thresholds, samples C10 and H5 contained fewer than 100 cells. C10 comprised tissue from only a single mouse, which could explain the low cell capture. Additionally, discarded cells in sample C10 highly upregulated myelin basic protein (*Mbp)*, suggesting that myelin removal was not sufficient (Müller et al., 2013), possibly effecting the capture efficiency of the sample. Supporting this was the fact that sample C10's infiltrating counterpart, B10, also had a very low yield (**Error! Reference source not found.**). T he other two 21 dpi infiltrating samples (G4 and F5.D21) had very low yields while their resident counterparts (F4 and E5.D21) were successful samples. At first, I assumed that there were fewer infiltrating cells at this time point, however, the IF-based quantification does not support this theory (Figure 4.5B), so it remains unclear why these samples were of low-quality. Ultimately, both C10 and H5 were removed as failed samples. In total, 23,174 cell libraries passed the QC filtering.

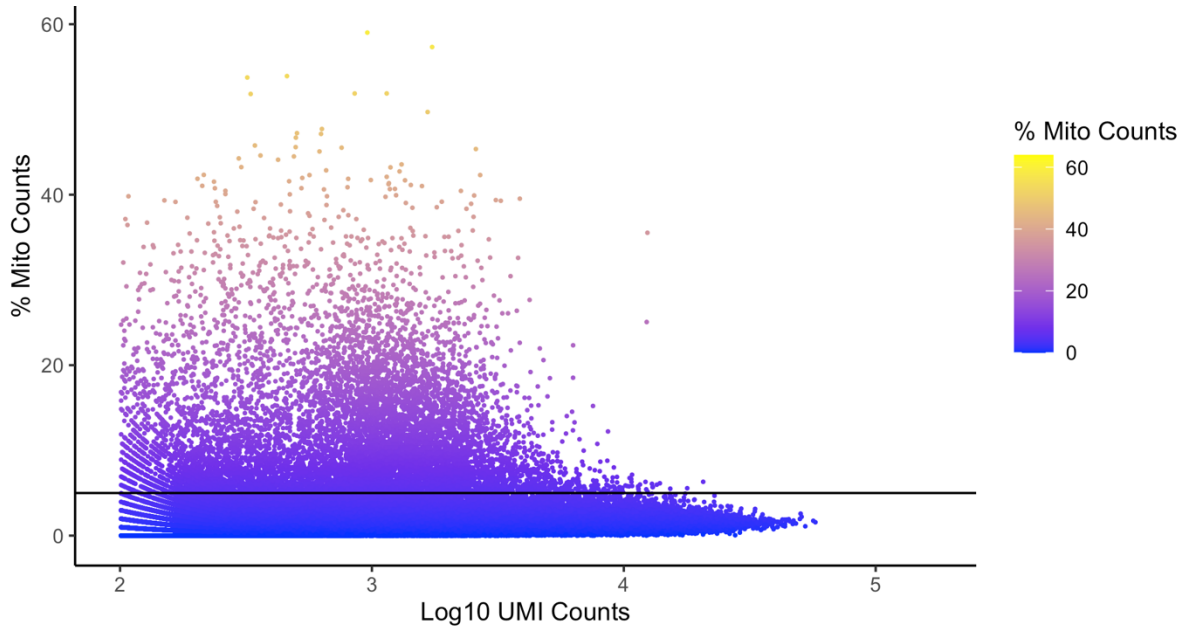


Figure 4.10 A plot of the log10 library size *vs* the proportion of mitochondrial transcripts for all the *CreMato* samples. Each cell is coloured by its proportion of mitochondrial transcripts. Cells with more than 5% mitochondrial counts were removed as low-quality.

As for the *Cx3cr1* dataset, I calculated the average log_2FC of each gene between the kept *vs* discarded cells and visualised these results by plotting the average UMI count of a given gene *vs* its log_2FC. Again, many of the increased genes in the discarded libraries were mitochondrial genes or non-myeloid CNS cells (Figure 4.11A). However, two notable differences were the strong presence of neutrophil genes in the retained libraries, and the increase of dendritic cell genes in the discarded libraries (Figure 4.11B). The neutrophil expression was driven by the 1 and 2 dpi infiltrating samples, in line with the reported temporal dynamics of neutrophil infiltration post-SCI (Donnelly & Popovich, 2008). Conversely, the increase in dendritic cell genes was driven only by sample B7 (Figure 4.11C, D). Given that sample B7 was dominated by cells with small libraries and low complexity, I chose not to alter the QC thresholds.

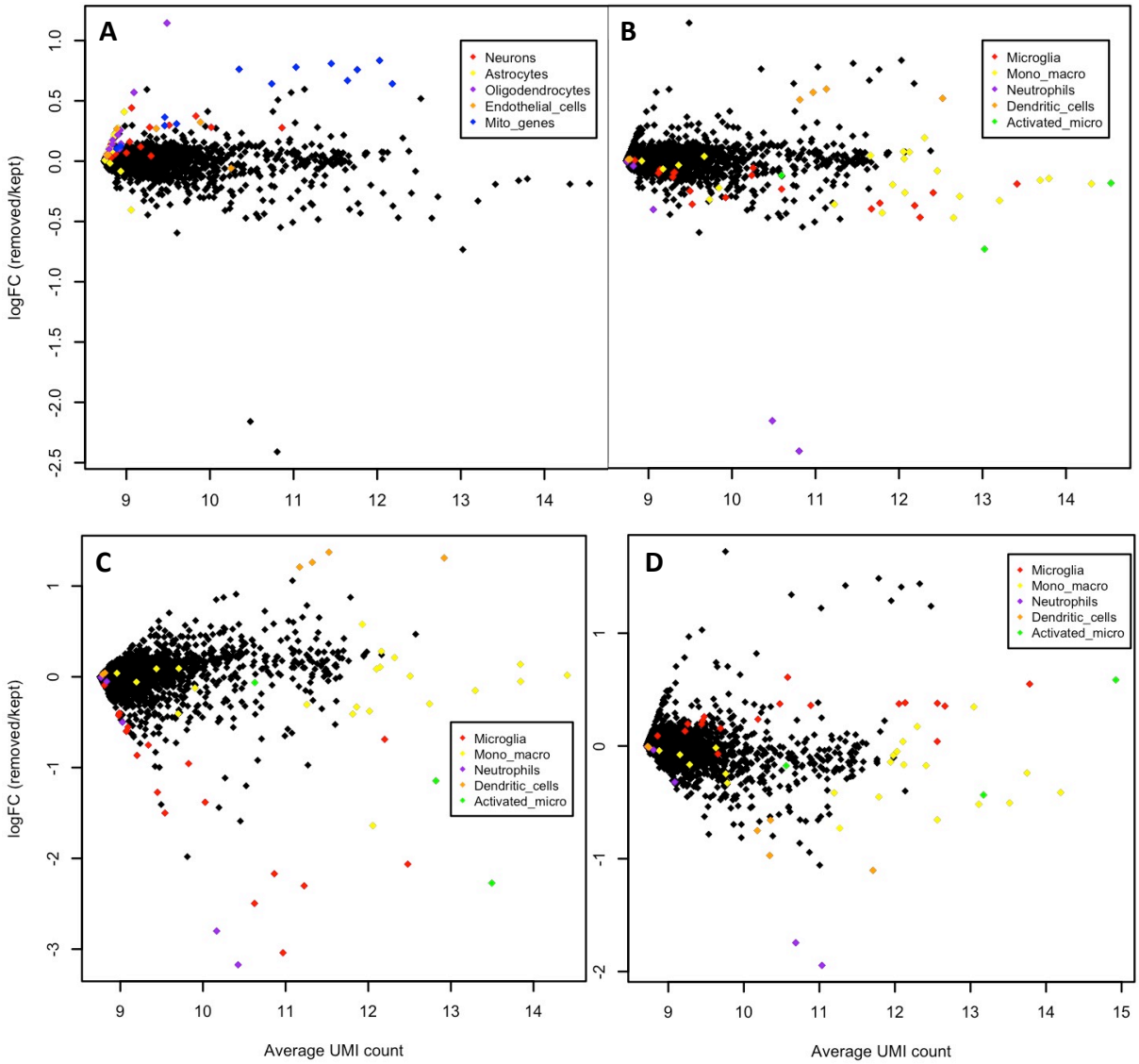


Figure 4.11 The average UMI count per gene *vs* the log_2 fold change between the cell libraries removed or kept based on QC thresholds. Each point is a gene. Genes are coloured by the cell type in which they are highly expressed. In general, expression of genes from non-myeloid CNS cells is higher in the discarded cell libraries (A) while myeloid cell genes are highly expressed in the cell that remain after QC, with the exception of dendritic cells which appeared to be prominent in the discarded libraries (B). However, it is clear that this effect was driven by sample B7 (C). If we re-examine the *CreMato* dataset without B7, dendritic cell genes, like other myeloid cell genes, are upregulated in the retained libraries.

4.5.3 Doublet Detection

To identify and remove doublets from the *CreMato* dataset, I called *cxds_bcds_hybrid()* function from the *scds* package, as in *3.4.5 Doublet Detection*. The hybrid output score predicted 2,464 doublets, which I removed from the dataset, leaving 20,710 cells for downstream processing (Table 4.1).

dpi	Sample ID	Fate map	Sex	# of mice	Collected	Flow Cell	Cell Count	Median Genes/Cell	Mean Reads/Cell
Ctrl	A10	resident	M	1	23-May-19	SLX-17850	176	738.5	6.07x104

						SLX-20619			
Ctrl	H2	resident	M	2	6-Mar-20	SLX-19290 SLX-20619	1,158	1,625.5	5.97x104
1	B11	infiltrating	M	3	31-Oct-19	SLX-18703	2,495	2,299.0	3.47x104
1	E5	resident	F	3	31-Oct-19	SLX-18703	1,580	2,774.5	4.41x104
2	F5	infiltrating	F	3	31-Oct-19	SLX-18703	1,627	1,580.0	4.55x104
2	A1	resident	M	3	6-Mar-20	SLX-19290	760	2,774.5	5.98x104
3	B7	infiltrating	M	2	26-Sept-19	SLX-18641	2,915	3,659.0	9.67x104
3	A7	resident	M	2	26-Sept-19	SLX-18641	777	918.0	5.64x104
3	D11	infiltrating	F	2	31-Oct-19	SLX-18703	3,579	2,831.0	5.37x104
3	C11	resident	F	2	31-Oct-19	SLX-18703	183	1,207.0	1.33x105
10	B10	infiltrating	M	1	23-May-19	SLX-17850	166	1,130.0	1.27x105
10	C10	resident	M	1	23-May-19	SLX-17850	removed	787.0	5.14x104
10	F3.D10	infiltrating	F	2	6-Mar-20	SLX-19290	521	2,025.0	9.89x104
10	E3.D10	resident	F	2	6-Mar-20	SLX-19290	2,731	1,212.0	9.42x104
21	G4	infiltrating	M	2	5-June-19	SLX-17998	158	1,143.5	3.88x105
21	F4	resident	M	2	5-June-19	SLX-17998	166	755.0	4.00x104
21	H5	infiltrating	M	1	1-May-21	SLX-20619	removed	755.0	5.31x104
21	G5	resident	M	1.	1-May-21	SLX-20619	616	1,588.5	6.69x104
21	F5.D21	infiltrating	F	2	1-May-21	SLX-20619	304	2,887.0	1.87x105
21	E5.D21	resident	F	2	1-May-21	SLX-20619	2,731	2,175.0	6.58x104

Table 4.1 The CreMato scRNAseq samples. M=male, F=Female. # of mice indicates the number of sex, age, and condition-matched mice pooled in the scRNAseq sample. The Cell Count, and Median Genes per cell all refer to the post-QC values. For the samples sequenced twice to increase sequencing depth, both flow cells are reported and the cell count, median genes/cell, and mean reads/cell are the results of the two combined sequencing runs.

4.5.4 CELL TYPE ANNOTATION

Given the *CellAssign* method of annotating cell types requires raw UMI counts, rather than log-normalised counts, I decided to annotate the cell types in the *CreMato* dataset before integrating it with the *Cx3cr1* atlas. This approach served as a 'sanity check', ensuring that the appropriate cell types were present and being corrected together across mouse strains.

With an 80% probability threshold, 1,929 cells were still unassigned and, in some cases, contradicted the fate mapping label. First, I manually annotated CAMs, defining them as any resident cell with ≥ 1 UMI count from *Cd74, Ms4a7,* and *Mrc1,* respectively. This definition corrected 22 dendritic cell labels and 25 microglia labels. Since both *CellAssign* and my manual annotations resulted in very few CAMs (0.3% of cells), I assumed that any unassigned resident cells were activated microglia, and manually annotated them as such**.** *Siglech* has been described as able to differentiate microglia from CAMs (Konishi et al., 2017), so I defined any resident cell with ≥ 1 *Siglech* count as microglia. For the unassigned infiltrating cells, I performed clustering as described in *3.5.1 Clustering* then assigned cell type based on the majority cell type of that cluster. Given that infiltrating macrophages and activated microglia can be particularly difficult to discern based on transcriptional profiles, I also used the fate

mapping labels to correct these two annotations, i.e., “infiltrating” microglia were re-annotated as macrophages, and vice versa.

4.6 Integrating and Comparing the Atlases

One of the main goals of this project is to compile and analyse a time-resolved myeloid cell atlas after SCI. Such analysis requires the integration of the *Cx3cr1* and *Cremato* datasets and the removal of technical heterogeneity, or batch effects. The first step in this process is multi-batch normalisation followed by HVG selection, and finally, batch correction. To begin this process, I harmonised the row and column data of the two pre-normalisation *SingleCellExperiment* objects and then concatenated them. The combined atlas contained 32,482 cells.

4.6.1 Normalisation

I normalised the combined dataset as described in *3.4.6 Normalisation*. As before, I defined batches as the flow cell lane except for samples sequenced twice. For those samples, I defined batches as samples sequenced in the same combination of flow cells (e.g., samples A10, C10, and D10).

With the resultant logcounts, I investigated whether the *CreMato* Ctrl samples might express the same cell cycling observed in the Cx3r1 sample, G11.D2, but did not observe a similar population in the *CreMato* dataset (Figure 4.12A). To be sure that the removed cells from G11.D2 were truly outliers, I performed the normalisation and batch correction steps with the *Cx3cr1* data that contained the cycling G11.D2 cells and found that the removed SIGAG11.D21 cells did not share substantial overlap with other Ctrl cells (Figure 4.12B). Thus, these cells were not included in the downstream processing.

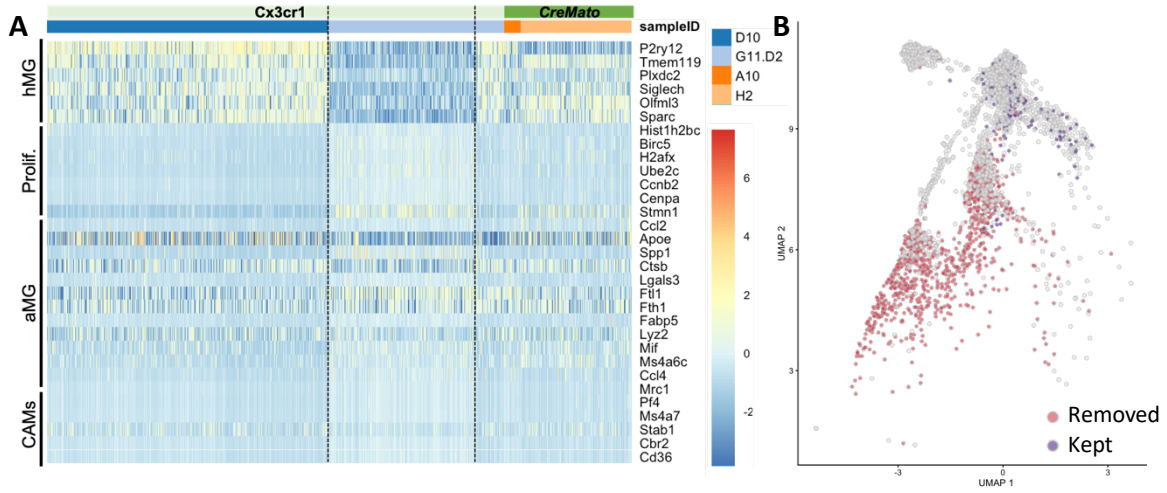


Figure 4.12 A) Heatmap of the Ctrl samples from the *Cx3cr1* and *Cremato* datasets. Each row is a gene, and each column is a cell. Genes are organised by the cell state they represent. hMG = homeostatic MG; Prolif. = proliferating MG; aMG = activated microglia; CAMs = CNS associated macrophages. Dotted lines highlight the cells from sample G11.D2 that upregulated proliferation markers and were removed from the *Cx3cr1* dataset. I did not observe a similarly dominant population in the *CreMato* dataset. B) A post-batch correction UMAP of the same control samples with sample G11.D2 in colour and other Ctrl cells in grey. G11.D2 cells that were removed do not generally overlap with other Ctrl samples.

As a quick 'sanity check' and final evidence that both the *Cx3cr1* and *CreMato* mouse lines were effective tools for isolating myeloid cells for scRNAseq, I also investigated the presence of non-myeloid cell markers in the data and did not observe any signs of non-myeloid cells (Figure 4.13) after QC.

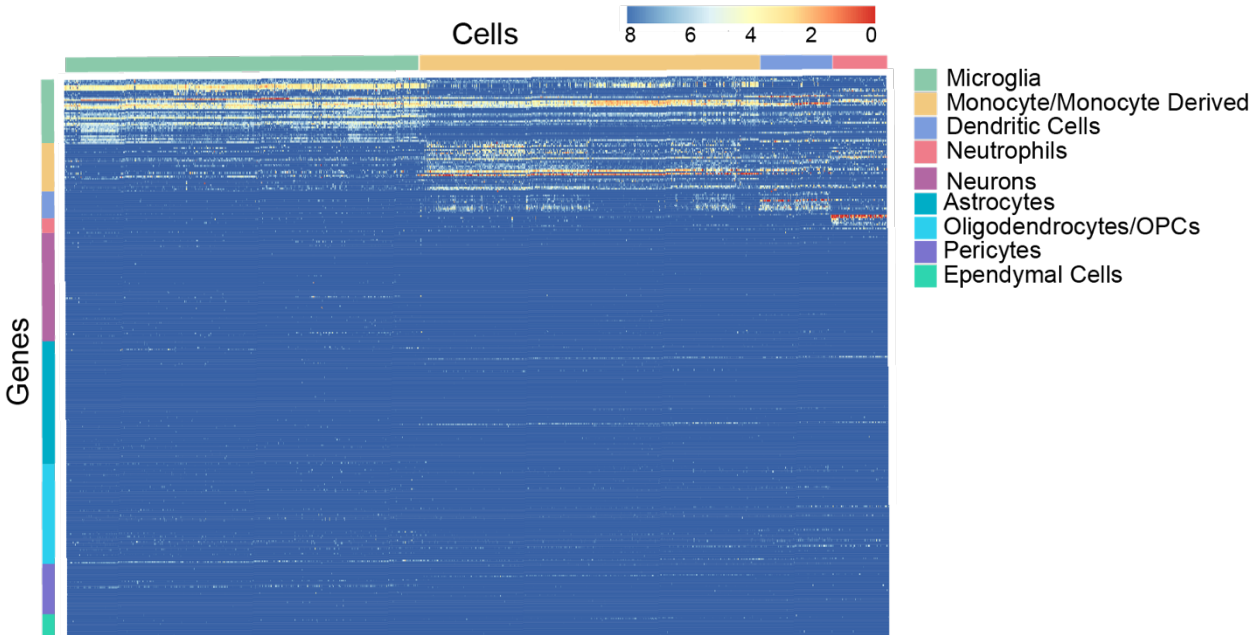


Figure 4.13 Heatmap of the scaled logcounts of canonical marker genes for CNS-associated cell types. I did not see any evidence of non-myeloid cell contamination in the integrated dataset.

4.6.2 Highly Variable Gene Selection

I selected HVGs as described in *3.4.6 Normalisation*, using a definition of batch that was consistent with the normalisation parameter. Unlike in the *Cx3cr1* dataset, the *CreMato* dataset contained low levels of T-cells and CAMs, ultimately increasing the biological heterogeneity of the dataset. So, I chose to increase the proportion of genes selected from those with biological variance from 20% to 40%. At 20% only 1 CAMs gene and 4 T-cell genes from **Error! Reference source not found.** were included in the HVG list, while at 40%, they w ere all included. Ultimately, a total of 3,131 HVGs were selected for the downstream analyses

4.6.3 Batch correction

Finally, I corrected for batch effects. Given the large number of collection dates (11) with far from perfect dpi distribution, I was reluctant to simply merge each collection date, in fear of inadvertently merging cells from subtly different stages in the SCI pathology. However, I had little information *a priori* as to which SCI stages would be identical, if any. So, I first visualised the pre-batch correction dataset by building a UMAP from the normalised logcounts and HVGs. I found that, without batch correction, separation was driven by the dpi, with relatively little separation between cell types collected at the same dpi, regardless of collection date (Figure 4.14A). In some cases, mouse strain also appeared to drive separation on the UMAP (Figure 4.14B). Additionally, as observed in the *Cx3cr1* dataset, the chemistry version produced a strong batch effect. Ultimately, I chose to batch correct based on strain and chemistry version, merging first the version 3 *Cx3cr1* and *Cremato* datasets, and then integrating the version 2 *Cx3cr1* batch. I performed this correction as described in *3.4.8 Batch Correction.*

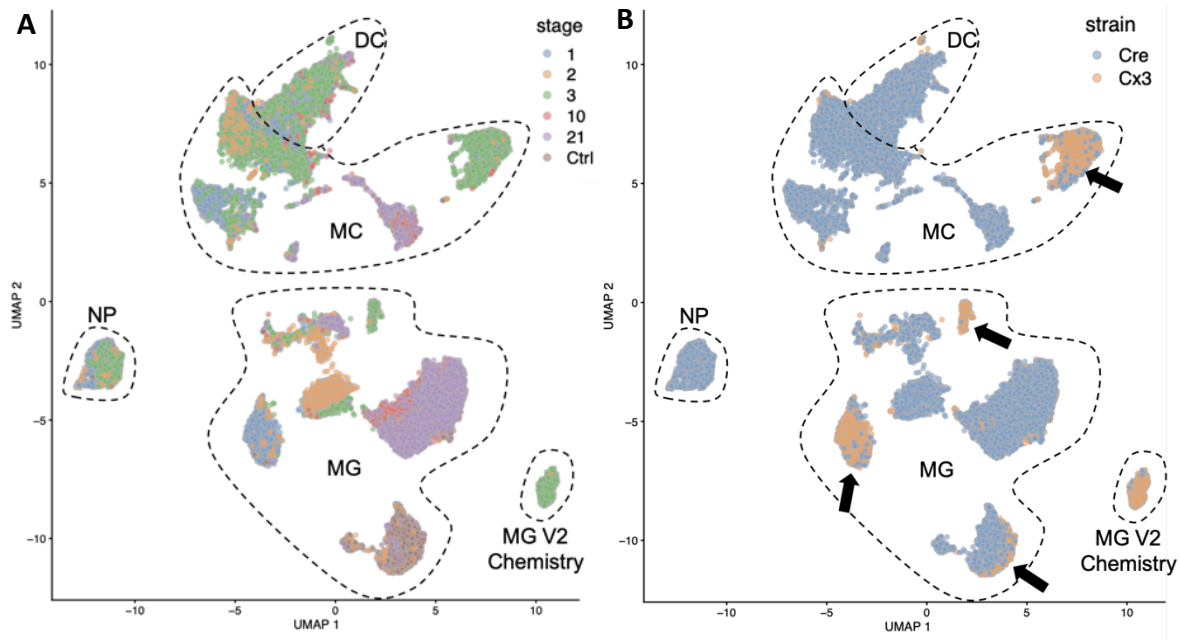


Figure 4.14 UMAPs of the combined dataset, before batch correction. A) Cells are coloured by the dpi. Dotted lines denote the majority cell type. Note that dpi and cell type are major drivers of separation in the UMAP, almost certainly demonstrating biological heterogeneity. MG = microglia, MC = monocytes & macrophages, DC = dendritic cells, NP = neutrophils. B) Cells are coloured by the strain of the mouse. Cre = *CreMato*, Cx3 = *Cx3cr1*. Arrows highlight discernible shifts between the strains. We can see that strain drives some heterogeneity, which, for the purposes of this study are uninteresting.

After batch correction, I evaluated the correction by visualising the dataset on a UMAP built from the batch-corrected coordinates. First, I noticed that cell types remained well-separated (Figure 4.15A), while overall, the strains were well-mixed (Figure 4.15B). I noticed that, as expected, dpi samples that overlapped before correction remained as such (e.g., 10 and 20 dpi microglia). Also, while the UMAP separation between cell types of the same dpi was reduced, it was not removed (e.g., 2 dpi *vs* 21 dpi microglia). Conversely, when I trialled batch correcting by collection date, there was substantial overlap between microglia, regardless of the dpi (Figure 4.15C). Finally, I investigated the variance lost within each batch during correction and found that while merging by collection date resulted in ≤ 16% of variance being lost, merging by strain and version removed only ≤ 2% of variance.

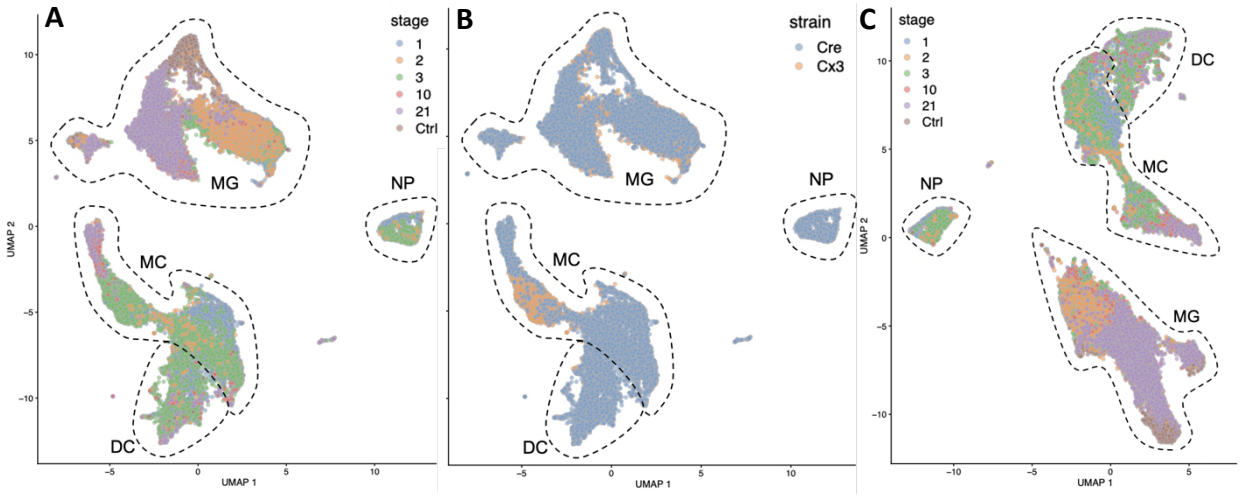


Figure 4.15 UMAPs of the combined dataset, after batch correction. A) Cells are coloured by the dpi. Dotted lines denote the majority cell type. Note that dpi and cell type still distinct in the UMAP, almost certainly demonstrating biological heterogeneity. MG = microglia, MC = monocytes & macrophages, DC = dendritic cells, NP = neutrophils. B) Cells are coloured by the strain of the mouse. Cre = *CreMato*, Cx3 = *Cx3cr1*. Arrows highlight discernible shifts between the strains. We can see that strain drives less heterogeneity after batch correction. C) UMAP of the dataset when batch correction was performed by collection date. Note that different SCI stages are merged too.

4.7 Data Sharing

Living in an era of big data, it is no longer feasible to explore and validate every aspect of a large dataset, such as this myeloid cell atlas. Rather than hiding the data away, with nothing but one's own biological question(s) addressed, there are plenty of avenues to making the data publicly available so that ultimately, the scientific community can get the most of out of datasets like this one.

For this dataset, I chose to make the data available through two common methods. First, I submitted the FASTQ files as well as the raw and pre-processed feature-barcode matrices to the publicly accessible Gene Expression Omnibus (GEO; GSE159638). Second, I built a Shiny-based interactive web app (https://marionilab.cruk.cam.ac.uk/SCI_Myeloid_Cell_Atlas/), allowing those without computational expertise to investigate gene expression patterns across different metrics, such as myeloid cell types and stages in the SCI pathology, through UMAPs and boxplots, or by downloading the pre-processed feature-barcode matrices (Figure 4.16). The code I used to build the app can be found here: https://github.com/regan-hamel/SCI_2020.

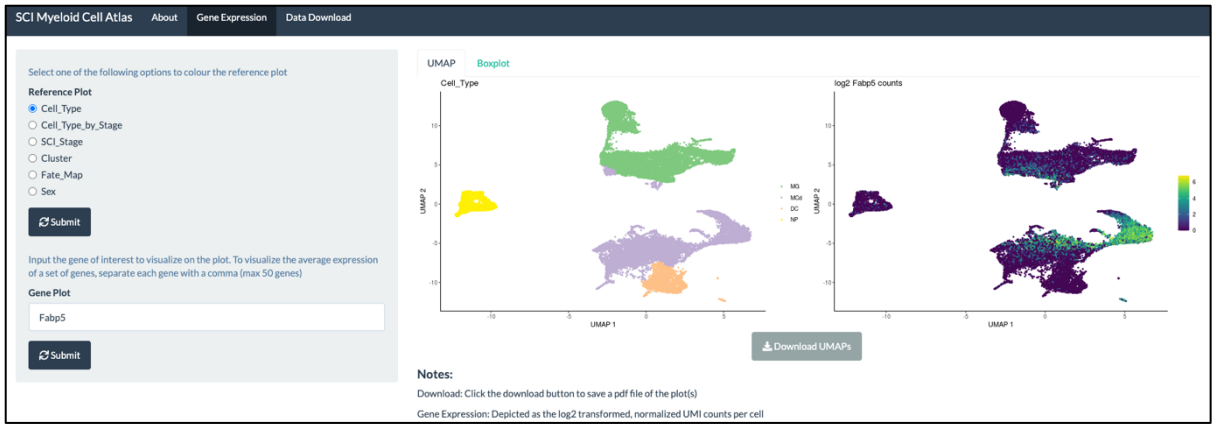


Figure 4.16 A quick look at the Shiny-based web app where researchers can easily explore the SCI myeloid cell atlas. The web app has three main pages denoted by the tabs in the top right corner. *About* is the landing page with a basic description of the project and experimental design, as well as useful links to websites including the relevant GitHub repositories. *Gene Expression* is the page depicted in the figure, where researchers can explore a gene of interest through UMAPs and boxplots. *Data Download*, as the name suggests, allows researchers to download the feature-barcode matrices for the entire dataset, or by cell type.

4.8 Discussion

In this chapter, I introduced a fate mapping *CreMato* mouse model and described the methodology I used to isolated infiltrating *vs* resident myeloid cells for scRNAseq. Many of the protocols and workflows were consistent with *Chapter 3*, with a few notable differences to accommodate for the new mouse strain. I also explained the methods I used to integrate the *CreMato* and *Cx3cr1* datasets in preparation for sharing these datasets with the wider research community and for my own analysis and interpretation in *Chapter 5*.

The *CreMato* mouse model offers several advantages over its *Cx3cr1* counterpart. Pertinently, it allows for the identification and separation of infiltrating *vs* resident myeloid cells through the expression of differentially expressed fluorescent reporter genes. This is an invaluable tool for discerning between the otherwise ambiguous activated microglia and infiltrating macrophages after SCI. Additionally, by enriching for these two populations separately, I was able to collect and ultimately characterise larger quantities of less well-represented populations such as dendritic cells, and even detect small quantities of rare populations like infiltrated CAMs.

Of course, the *CreMato* model is not without its caveats. As considered in *3.6 Discussion,* enriching for cell populations of interest before scRNAseq can limit the types of analyses that can be performed on the resulting dataset. Additionally, fluorescent reporter mice, while intended to have very specific expression patterns, often suffer from "*leakiness,*" or

unintended expression patterns. In the case of the *CreMato* mouse, we had to account for the "*leakiness*" of two fluorescent lines.

Firstly, as mentioned in *3.6 Discussion*, the *Cx3cr1* mouse line (Parkhurst et al., 2013) has been reported to leak into neurons (Zhang et al., 2018; Zhao et al., 2019), possibly due to their transient expression of *Cx3cr1* (Dworzak et al., 2015; Wang et al., 2018). Furthermore, the *Cx3cr1* mouse line was generated via BAC transgenics, which retains the function of the gene of interest, but can actually lead to duplication of said gene (Zhao et al., 2019). It is recommended to maintain such transgenic lines as heterozygotes, however, in the case of the *CreMato* mouse, this is not practical as it would result in three quarters of offspring being unusable. Thus, we risked unexpected effects from *Cx3cr1* over-expression. To account for this, we can compare our findings with studies that employed other mouse strains, although this approach only works for unoriginal results.

Secondly, we have the tdTomato^{flox} line, which has been reported to, through an unknown mechanism, have variable leakage into neurons, Iba1^{+} cells, and astrocytes (Zhao et al., 2019). In the case of the non-tamoxifen treated *CreMato* mouse, this leakage was observed in CAMs and microglia, but not monocytes, suggesting that the line remains a valid tool for distinguishing between infiltrating and resident myeloid cells (Jordão et al., 2019).

At later stages in the SCI pathology, I noticed that the intensity of YFP in resident myeloid cells varied. A possible explanation for this could stem from the dynamics of the fluorescent proteins. The *CreMato* mouse expresses YFP in *Cx3cr1*$^{+}$ cells, which, under homeostatic conditions, means all myeloid cells produce a YFP transcript each time *Cx3cr1* is transcribed. When microglia become activated, they downregulate homeostatic marker genes, such as *Cx3cr1* (Dubbelaar et al., 2018). Given the long half-life of YFP (> 24hrs (Hentschel et al., 2013)) and the presumably low-level of continued *Cx3cr1* expression, YFP continues to be expressed in activated microglia. However, activated and proliferating microglia would dilute their cytosolic YFP and continue to express *Cx3cr1* at low levels, meaning less YFP would be present in those cells. The same would be true for tdTomato, which is also under the *Cx3cr1* promoter. However, tdTomato is 6x brighter than YFP and doesn't need to be enhanced before microscopy, which could explain why there was no obvious decrease in its intensity a 21 dpi

(Figure 4.6B). This range of YFP and tdTomato fluorescence has implications for the FACS isolation method used to prepare the scRNAseq samples. In an effort to isolate pure infiltrating *vs* resident myeloid cell populations, we were very strict with our gating strategy. However, we may have inadvertently enriched for highly YFP^{+} and highly $tdTomato^{+}$ myeloid cells.

To address this question, and others that will be discussed in *5.6 Discussion*, I designed a proliferation-tracking BrdU experiment to label acutely proliferating cells in *CreMato* tissue, which could then be stained for BrdU, YFP, tdTomato, and several mRNAs of interest using smFISH. Under this framework, I would be able to check whether YFP^{lo} microglia at 21 dpi were $BrdU^{+}$ and thus acutely proliferating. Unfortunately, at the time of submission I had performed the *in vivo* experiment and prepared the tissue but had not yet received the final data from Katherine.

To be able to ultimately integrate the *CreMato* atlas with the *Cx3cr1* data, I maintained the same protocols and computational workflows for the preparation of both datasets whenever feasible. During the QC phase, I noticed two major differences in the genes present in retained *vs* discarded libraries. First, I found that dendritic cell transcripts were increased in the discarded libraries and that this effect was driven by a single sample, B7. From the violin plots of library size and complexity distributions (Figure 4.8), it is clear that B7 contains a very large population of very small cell libraries with little complexity. This aberrant population was removed by the QC thresholds and, given that it was not present in other 3 dpi infiltrating or *Cx3cr1* samples, likely represented a technical artefact. The second notable difference was in neutrophil genes, which were increased in the retained libraries. Compared to the *Cx3cr1* dataset, the *CreMato* dataset isolated more neutrophils, as expected, and perhaps more neutrophils passed the QC thresholds compared to other myeloid cell types as their smaller size might make them less prone to damage during mechanical homogenisation.

Despite my best efforts (see *3.2 Myeloid Cell Isolation*), a major challenge in this study was collecting high quality samples with high cell counts, stemming from several constraints and mistakes. Given that we were interested in the dynamics of myeloid cells at the lesion epicentre, and all spatial context is lost during scRNAseq, we chose to dissect very small tissue

sections (5 mm) for sequencing. This meant that cell yields were typically very low (< 1000), unless several mice were pooled together. Unfortunately, pooling several sex and dpi matched littermates was not always feasible due to animal logistics, including project licence restrictions on the number of mice that can receive SCI, and the rate at which mice would reach their humane endpoint before their collection date. However, data from a recent scRNAseq of all CNS cells after SCI, which pooled 8 mm of tissue from 5 mice, only collected 2-3 times more cells per sample (Milich et al., 2021). Thus, the best way to overcome this would be to use a technology that retains spatial context, such as GeoMX WTA (Roberts et al., 2021). Another constraint was the use of mechanical homogenisation. This approach was necessary to minimise *ex vivo* activation, but also generated substantial ambient mRNA, and likely damaged many myeloid cells. Future studies requiring the *ex vivo* isolation of microglia should certainly employ the new transcription/translation-inhibited enzymatic dissociation protocol (Marsh et al., 2020). Finally, an error in the isolation protocol was the FACS-gating strategy that we adopted, which was likely too strict and excluded a proportion of myeloid cells. Given that myeloid and non-myeloid cells are easily discernible *in silico*, this should have been given less of priority over cell yield.

Despite these challenges, I was able to generate a time-resolved myeloid cell atlas of 32,482 high quality cells with several thousand cells per condition. This atlas presents an exciting opportunity to investigate the dynamics of myeloid cells after SCI, and will hopefully help future researchers to identify novel therapeutic approaches to combat the disability inflicted by SCI.

Chapter 5 CHARACTERISING THE TIME-RESOLVED MYELOID CELL ATLAS

"For a research worker the unforgotten moments of life are those rare ones which come after years of plodding work, when the veil over natures secret seems suddenly to lift & when what was dark & chaotic appears in a clear & beautiful light & pattern."

— Gerty Cori

DISCLOSURE

The *in silico* analysis was designed and performed by me. Veronica Testa, a visiting master's student from the University of Milan Bicocca, helped me with the tissue preparation for IF and smFISH. The IF and smFISH was accomplished in collaboration, as indicated in the relevant sections.

Certain aspects of this chapter are expanded from the bioRxiv pre-print: Hamel, R. *et al.* Time-resolved single-cell RNAseq profiling identifies a novel Fabp5-expressing subpopulation of inflammatory myeloid cells in chronic spinal cord injury. *bioRxiv*, 2020.2010.2021.346635, doi:10.1101/2020.10.21.346635 (2020). I wrote the manuscript for this pre-print in its entirety. Luca Peruzzotti-Jametti, Stefano Pluchino, and John Marioni edited the manuscript.

The *Fabp5* discussion draws from the *Lipid Metabolism* section of the review: Peruzzotti-Jametti, L., Willis, C. M., Hamel, R., Krzak, G. & Pluchino, S. Metabolic Control of Smoldering Neuroinflammation. *Frontiers in Immunology* 12, doi:10.3389/fimmu.2021.705920 (2021). I wrote this section, and it was edited by Luca Peruzzotti-Jametti.

5.1 Introduction

scRNAseq data requires substantial low-level analysis as described in Chapters 3 & 4. However, the ultimate purpose of generating this type of data is to make new biological discoveries at single cell resolution. Approaches to facilitate biological interpretation and discovery through scRNAseq are flexible and numerous. In this chapter, I first verify that the novel dataset contains the expected cell types and states, the markers of which have become increasingly available in the literature over the course of this PhD. After establishing this, I take the typical first step in exploratory analysis: grouping transcriptionally similar cells through unsupervised clustering. Then, I investigate the cluster dynamics after SCI through trajectory analysis. From there, I characterise the clusters of interest, drawing heavily from previous findings and resources such as gene ontology (GO) and PanglaoDB, and finally present a summary of proposed myeloid cell dynamics after SCI. The code I wrote to handling the data and generate the figures for this chapter, including the smFISH analysis I performed, are publicly available on my GitHub (https://github.com/regan-hamel/PhD_Dissertation).

5.2 Examining the Myeloid Cell Atlas for Known Phenotypes

5.2.1 The Dynamics of Marker Gene Expression Upon Activation

Since beginning this PhD project, many studies have described myeloid cell heterogeneity at single cell resolution across development, homeostasis, and a wide range of pathologies. A common observation across several CNS pathologies, including also very recently SCI, is the downregulation of canonical microglia markers (Bennett et al., 2018; Butovsky et al., 2014; Dubbelaar et al., 2018; Hammond et al., 2019; Jordão et al., 2019; Milich et al., 2021). Given this, I investigated the expression of these genes in microglia over time and did indeed observe a downregulation in *Serpine2*, *Tmem119*, *P2ry12, Cx3cr1, Siglech, Plxdc2, P2ry12,* and *Gpr34* (Figure 5.1A). These genes appeared to be exclusively expressed in microglia in the acute phase, but by the subacute (10 dpi) and early chronic (21 dpi) phases, macrophages upregulated these genes, with the exception of *Siglech* (Figure 5.1B) (Konishi et al., 2017). *Serpine2* (glia-derived nexin) also appeared to be exclusive to microglia at later time points and, unlike *Siglech*, based on the scRNAseq data it appeared to be downregulated only at 1 dpi. Other microglia genes, including *Hexb, Csf1r, Ctss, C1q, Trem2, Sparc, Olfml3,* and *Fcrls,*

remained highly expressed in microglia after SCI but were upregulated in acute infiltrating myeloid cells (Figure 5.1B).

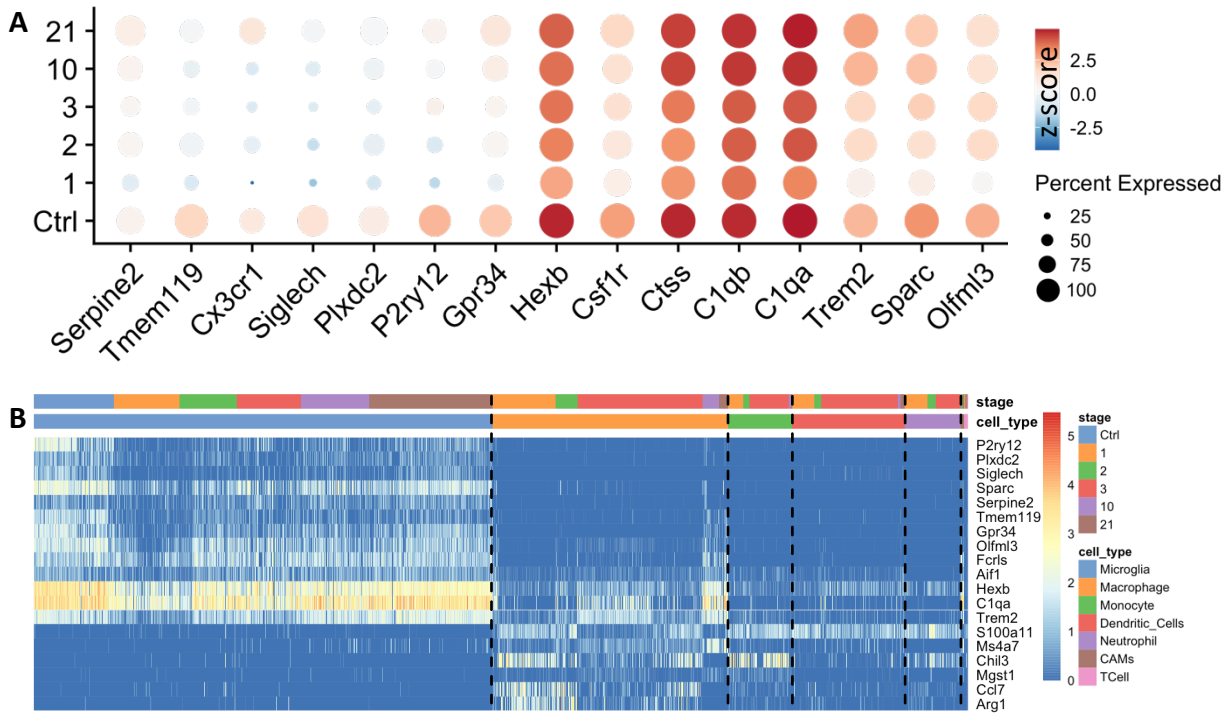


Figure 5.1 A) Dot plot showing canonical microglia marker expression across the dpi time points. B) A heatmap of the expression of myeloid cell markers across cell types and different SCI conditions. Each row is a gene, and each column is a cell. Dotted lines highlight breaks between cell types.

I investigated typical markers of infiltrating myeloid cells and found that *Chil3* (Ym1) was expressed by monocytes, macrophages, and neutrophils (Gensel & Zhang, 2015), and *Ms4a7* (Bennett et al., 2018), *Ccl7* (Zhu et al., 2017), *Arg1* (Jordão et al., 2019), and, to a lesser extent *Mgst1* (Haage et al., 2019), were restricted specifically to monocytes and macrophages (Figure 5.1B). Notably, at 10 and 21 dpi, only *Ms4a7* remained highly expressed in macrophages and CAMs (Figure 5.2A). Also of note was the expression pattern of *S100a11*, which appeared to be restricted to infiltrating cells across all post-SCI time points (Figure 5.2B).

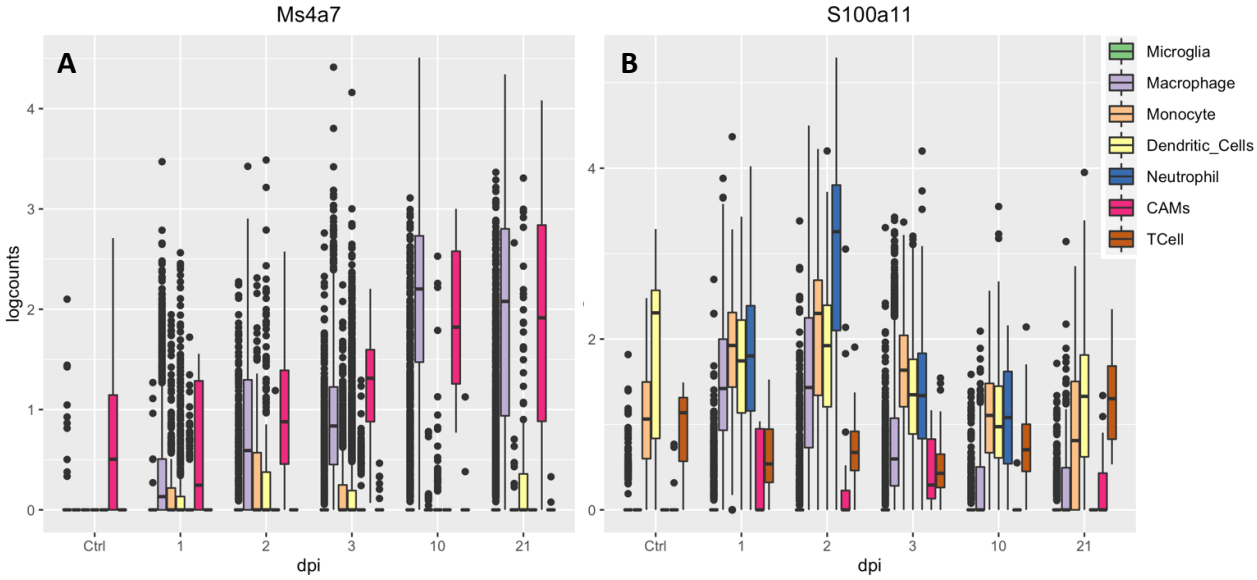


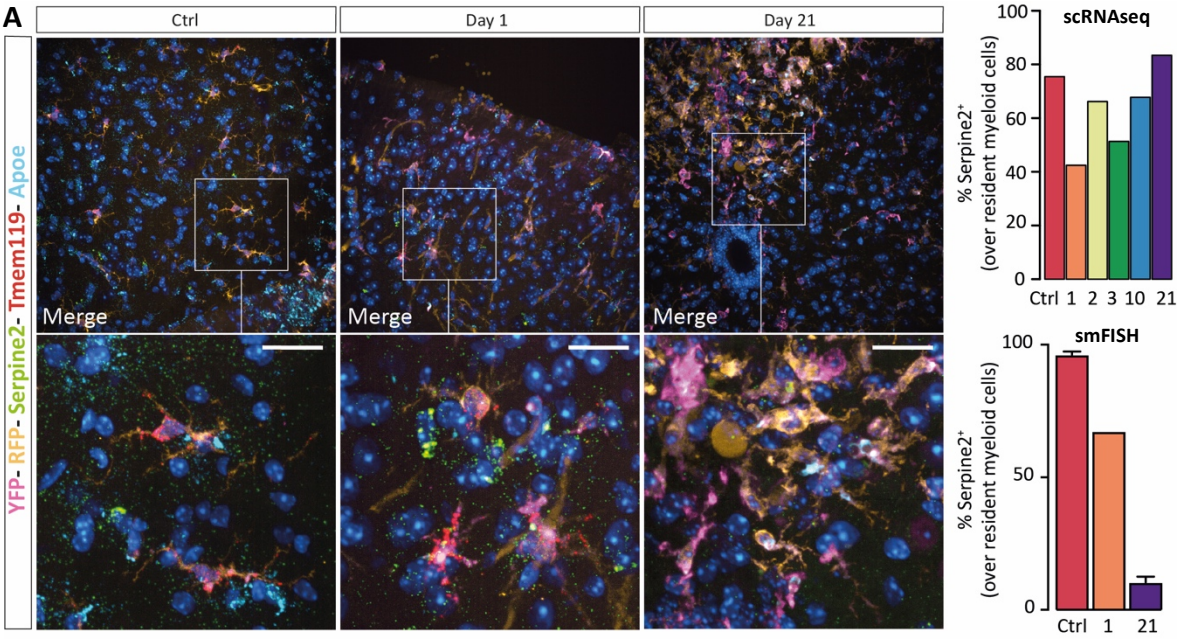
Figure 5.2 Boxplots by cell type and time point demonstrating A) the expression of *Ms4a7* in macrophages and CAMs, persisting into the chronic phase and B) the restricted expression of *S100a11* to infiltrating myeloid cells. Bins in each boxplot are ordered from left to right: microglia, macrophages, monocytes, dendritic cells, neutrophils, CAMs, and T-Cells.

To confirm the dynamic expression of *Serpine2* and *Ms4a7*, we probed their expression in *CreMato* tissue via smFISH. For the smFISH quantification, cells with a spot count > 5 were considered positive for that gene. *Tmem119, Apoe,* and *S100a11* were also probed, but not quantified. As described in *4.2.2*, Veronica and I prepared the *CreMato* tissue for smFISH, which was planned by Stefano, Katherine, and I. Katherine performed the staining, imaging, and analysis. The expression of *Serpine2* was not exclusive to myeloid cells, in line with previous reports of its expression in astrocytes (Y. Yang et al., 2018) (Figure 5.3A). We found that Ctrl microglia expressed *Tmem119* and *Serpine2*, but not *Apoe* as expected, although *Apoe* was present in the Ctrl tissue (Figure 5.3A), in line with reports of *Apoe* expression in homeostatic astrocytes (Gee & Keller, 2005). The proportion of *Serpine2*$^{+}$ resident myeloid cells decreased by 1 dpi, however, at 21 dpi, they remained low according to the smFISH data, while the scRNAseq data suggested a return to homeostatic proportions. This could be due to the different anatomical regions observed – the scRNAseq data was generated from 5 mm of tissue including normal appearing spinal cord (NASC), while the smFISH data is restricted to the lesion epicentre. Other potentially confounding factors will be further addressed in the *Discussion*.

The proportion of *Ms4a7*$^{+}$ infiltrating cells was fairly consistent between the scRNAseq and smFISH data in the Ctrl and acute stages of the pathology – low in the Ctrl and increasing to 3

dpi, albeit with more *Ms4a7*$^{+}$ cells detected at 1 dpi in the scRNAseq data (Figure 5.3B). However, by 21 dpi, the proportion of *Ms4a7*$^{+}$ infiltrating cells was much higher in the scRNAseq data than the smFISH. This may again reflect lesion-specific phenotypes.

S100a11 did not appear to be expressed in resident myeloid cells, but was not restricted to infiltrating myeloid cells, in line with reports of its expression in endothelial cells (Franzén et al., 2019). Thus, we confirmed the presence of several genes of interest in the tissue but were unable to perfectly recapitulate the expression patterns observed in the scRNAseq data. Given the difference in anatomical regions, this does not necessarily negate the accuracy of the scRNAseq data, but neither does it fully validate it.



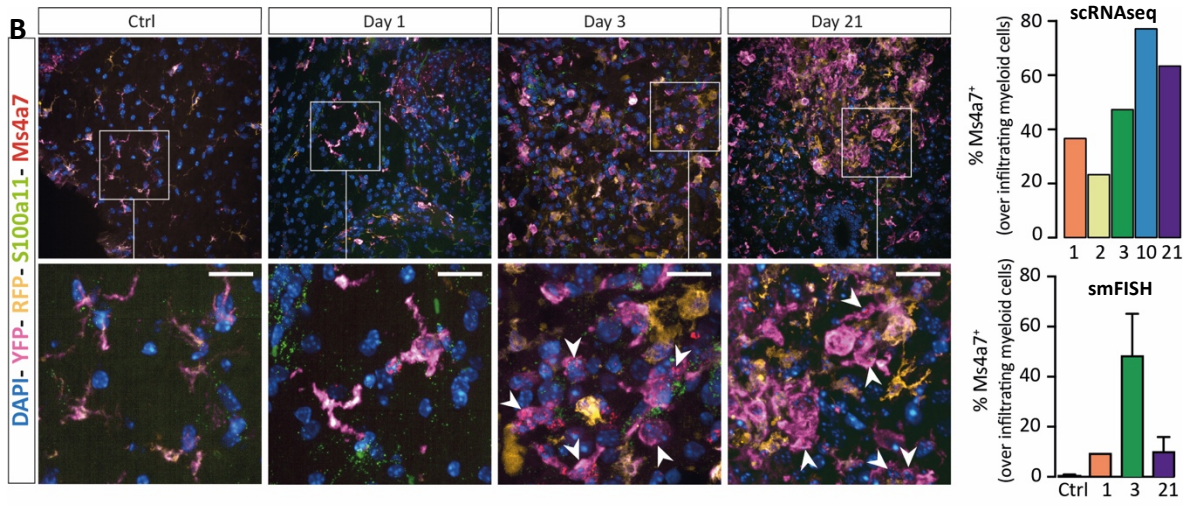


Figure 5.3 A) smFISH depicting the expression of *Tmem119*, Serpine2 and *Apoe* from Ctrl, 1, and 21 dpi. The bar graph shows the quantification of *Serpine2*$^+$ cells out of all RFP$^+$/YFP$^+$ resident myeloid cells from scRNAseq (top) and smFISH (bottom). B) Expression of *Ms4a7* and *S100a11* from Ctrl, 1, 3 and 21 dpi. The bar graphs show the quantification of *Ms4a7*$^+$ cells over or all RFP$^-$/YFP$^+$ infiltrating myeloid cells from scRNAseq (top) and smFISH (bottom). Arrowheads indicate *Ms4a7*$^+$/RFP$^-$/YFP$^+$ cells. smFISH data are mean % (± SD) from n= 2 mice per time point. Nuclei were stained with DAPI. Scale bars: 20 µm.

5.2.2 The TREM2-APOE Pathway

Several studies of myeloid cells under neuropathological conditions have observed the stark upregulation of *Apoe* upon myeloid cell activation (Iwata et al., 2005; Kim et al., 2015; Shin et al., 2014), including after SCI (Seitz et al., 2003; X. Yang et al., 2018). To confirm this pattern in this myeloid cell atlas, I calculated the average expression of each gene in the Ctrl and after SCI and plotted the top 15 most highly expressed genes for each condition (Figure 5.4). While the Ctrl was dominated by homeostatic microglia markers as expected, *Apoe* was the most highly expressed gene in the SCI samples, with a strong upregulation in microglia and macrophages from 3 dpi.

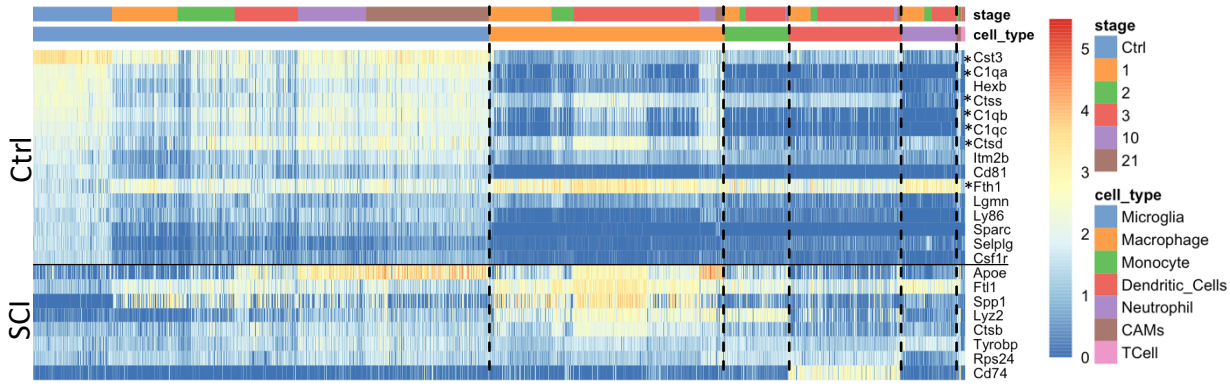


Figure 5.4 A heatmap of the expression of the 15 most highly expressed genes in Ctrl and SCI myeloid cells across cell types and different SCI conditions. Lists are separated by solid black line, * indicates duplicates between the two lists. Each row is a gene, and each column is a cell. Dotted lines highlight breaks between cell types.

As described in *1.3.4*, recently, the expression of *Apoe* and the activation of the TREM2-APOE pathway has been shown to mediate the switch from homeostatic to activated microglia by sensing neurodegeneration-associated molecular patterns, including myelin debris and apoptotic neural cell bodies (Deczkowska et al., 2018; Keren-Shaul et al., 2017; Krasemann et al., 2017). The upregulation of the TREM2 adaptors, *Apoe* and *Tyrobp*, and the downregulation of homeostatic microglia markers, are described as the first of two DAM phases (Friedman et al., 2018; Keren-Shaul et al., 2017). The DAM2 phase is characterised by the upregulation of phagocytosis, lysosomal, and lipid metabolism genes and has been described as a neuroprotective, phagocytic, phenotype in mouse models of AD, ALS, MS and ageing (Deczkowska et al., 2018; Keren-Shaul et al., 2017; Krasemann et al., 2017). Notably, the DAM phenotype appears to be conserved between mice and humans (Keren-Shaul et al., 2017), and has been suggested to be a disease-independent microglia response (Deczkowska et al., 2018). So, although the DAM transcriptional profile has yet to be described in SCI, I expected to find these phenotypes in the myeloid cell atlas.

I investigated the expression of the homeostatic, DAM1, and DAM2 transcriptional profiles in microglia from the myeloid cell atlas (Figure 5.5). Microglia downregulated the expected homeostatic markers and by 2 dpi began to adopt DAM1-like phenotypes. DAM2-like phenotypes were present by 10 dpi, although *Axl*, *Csf1*, and *Itgax* were only slightly upregulated and *Lilrb4a* was not in the HVGs, suggesting a degree of disease specificity in the DAM2 phenotype. By 21 dpi, despite the strong upregulation of DAM1 and DAM2-associated genes, many, but not all, microglia began to recover the expression of homoeostatic genes, including *Cx3cr1.*

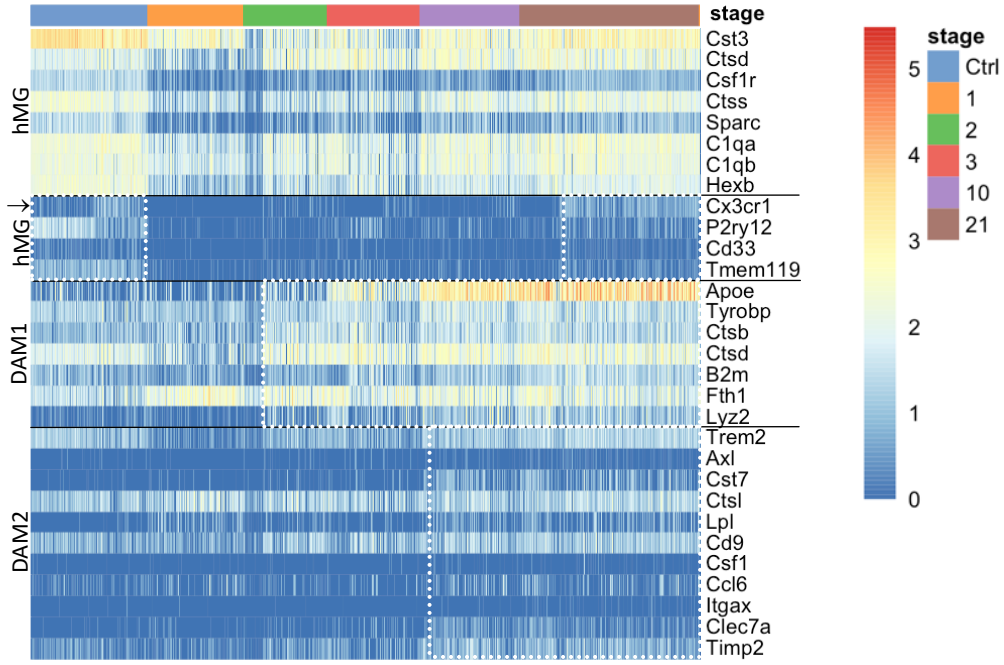


Figure 5.5 A heatmap of homeostatic and DAM-associated genes in microglia across the different SCI conditions. Profiles are separated by solid black line, hMG = homeostatic microglia, hMG↓ = homeostatic microglia that are downregulated in the DAM1 signature. Each row is a gene, and each column is a cell. Dotted white boxes highlight microglia that fit the designated profile.

5.2.3 Microglial Proliferation Post-SCI

Based on previous reports, I expected a proportion of the myeloid cell atlas to be proliferating in response to the SCI (Bellver-Landete et al., 2019; Greenhalgh & David, 2014; Milich et al., 2021; Noristani et al., 2017; Wahane et al., 2021). To identify and annotate these cells, I called the *cyclone* function from the *scran* package on a *SingleCellExperiment* object with the full list of genes (*vs* just HVGs) and used the package's pre-trained mouse marker gene pairs (Scialdone et al., 2015). Briefly, this method classifies cells of unknown cell cycle phase by investigating the expression of pre-identified pairs of genes that change their relative expression in a predictable direction according to the cell cycle phase. This approach was able to annotate cells as G1, S, or G2M, but over half the cells remained unassigned. Based on a published scRNAseq microglia cell cycle analysis, I manually annotated the unassigned cells as G0 (Li et al., 2019).

To corroborate the cycling annotations, I investigated the expression of microglia proliferation profiles (Hammond et al., 2019; Milich et al., 2021) across the cycling phase annotations (Figure 5.6). While many clearly cycling cells were annotated as such, especially in the G2M phase, there was evidence of a population of cycling microglia, predominantly from Ctrl samples, that were unassigned and consequently annotated as G0. As I had no way

of determining whether these G0 cells were truly cycling, I decided not to re-annotate them, and instead used the G0 phase annotations with caution when inferring biological relevance.

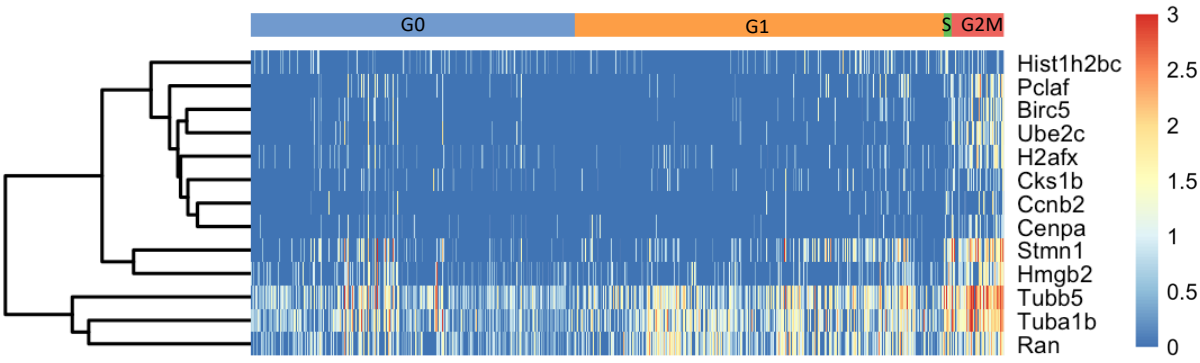


Figure 5.6 Heatmap of microglia from the myeloid cell atlas and their expression of proliferating microglia-associated genes. Each row is a gene, and each column is a cell. Genes are organised by the annotated cell cycling phase. G1 phase cells slightly upregulate proliferating genes, while G2M phase genes closely match the proliferation profiles. Some G0 cells upregulate these markers and may have been mislabelled.

Knowing that microglia undergo rapid proliferation in the acute phase of SCI (Bellver-Landete et al., 2019), I used an area plot to visualise the cell cycling phase over time post-SCI (Figure 5.7A) and saw that, as expected, G1 cells peaked at 1-2 dpi, and S/G2M peaked around 2-3 dpi, with cycling cells decreasing over time, in line with the findings of recent scRNAseq studies of microglia after SCI (Milich et al., 2021; Wahane et al., 2021). I further visualised these annotations by projecting them onto a UMAP of the myeloid cell atlas SCI (Figure 5.7B) and observed that cycling microglia from 1-2 dpi formed a large visual cluster on the UMAP, supporting the notion that the *cyclone* annotations were biologically relevant.

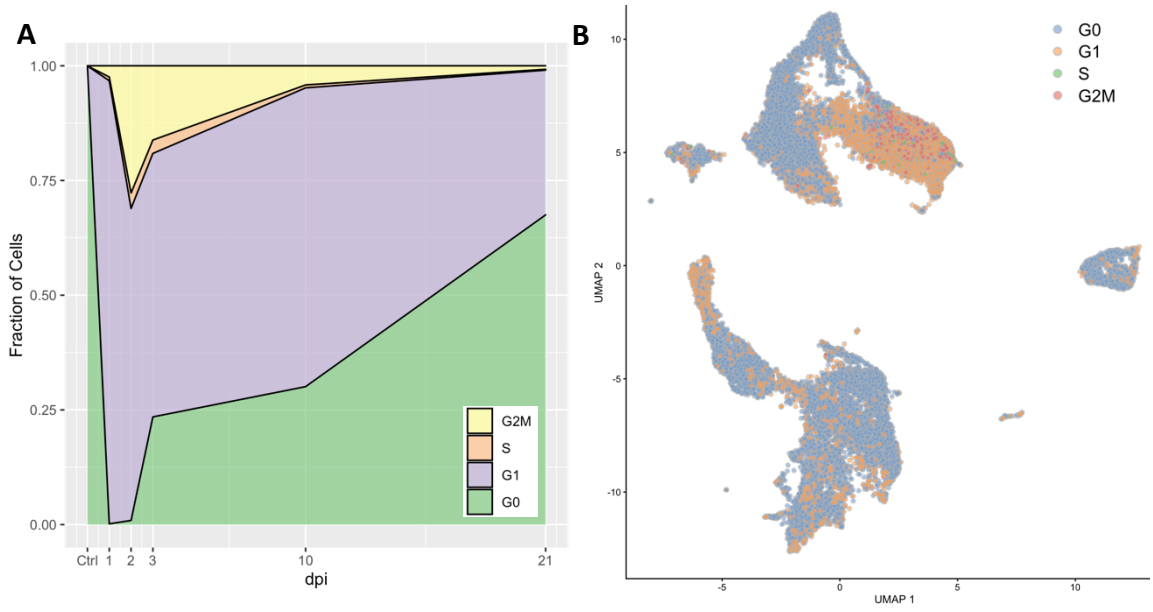


Figure 5.7 A) Area composition plot of the cell cycle of microglia captured at each stage of SCI. The dynamics are in line with previous reports. B) UMAP of the myeloid cell atlas coloured by cell cycle phase. Microglia from 1-2 dpi are overwhelmingly cycling. A population of macrophages from 10 and 21 dpi also appear to be cycling.

A recent study that isolated microglia nuclei using the INTACT method observed robust upregulated of *Hdac3* in microglia in mouse models of thoracic transection SCI (Wahane et al., 2021). This gene was hypothesised to control the proliferative state of microglia after SCI. In this myeloid cell atlas, *Hdac3* expression was minimal and was not included within the HVGs.

5.2.4 M1/M2 Polarisation

Over the past few years, the relevance of the M1/M2 polarisation *in/ex vivo* has faded, partially driven by the fact that many scRNAseq studies have not found evidence to support the M1/M2 phenotypes in activated myeloid cells (Kim et al., 2016; Milich et al., 2021; Ransohoff, 2016). Despite this, for completeness, I investigated these genes upon SCI and did not observe a clear separation (Figure 5.8). In fact, M1 and M2 genes were generally very lowly expressed, or even absent from the HVGs (e.g., *Il6*, *Il12a*, *Pparg*), and many activated myeloid cells expressed both M1 and M2 genes.

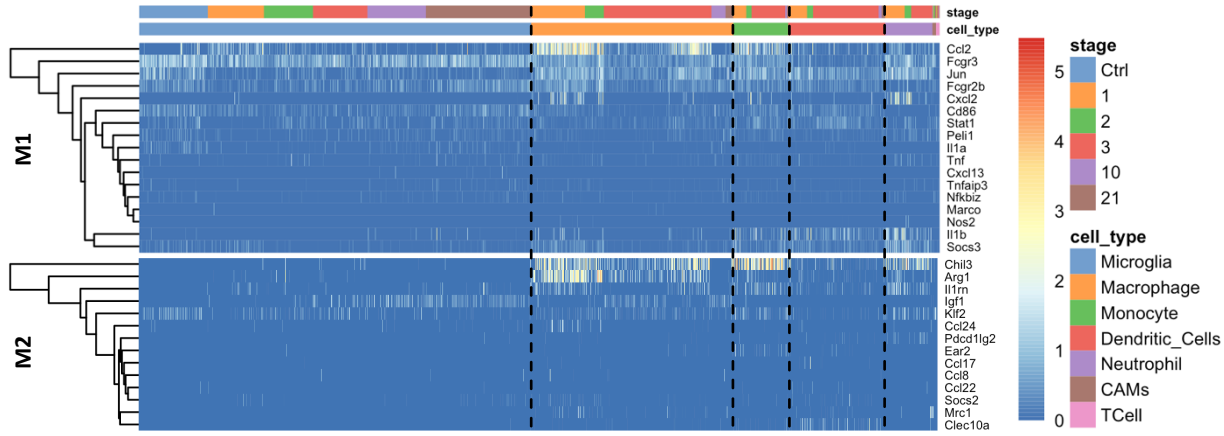


Figure 5.8 A heatmap of M1 and M2 markers in myeloid cells across cell types and different SCI conditions. Each row is a gene, and each column is a cell. Dotted lines highlight breaks between cell types. There was no clear match for M1 or M2 profiles to the myeloid cell atlas.

5.2.5 Sexual Dimorphism in Microglia

Previous studies have reported conflicting evidence for sexual dimorphism in microglia, especially during development and disease (Masuda et al., 2020). Early scRNAseq studies have observed minimal, if any, differences between sexes (Hammond et al., 2019; Tay et al., 2018). Several more recent studies have suggested there are indeed sex-specific roles in the microglia response to disease pathology (Han et al., 2021; Ochocka et al., 2021; Rahimian et

al., 2019). Unfortunately, the time points across male and female samples were not perfectly matched (e.g., there is no 1 dpi *Cremato* resident female sample). Thus, it appeared that there were shifts between sexes (e.g., 1 & 3 dpi infiltrating cells), but I could not confidently draw any sexual dimorphism-related conclusion conclusions from this dataset (Figure 5.9).

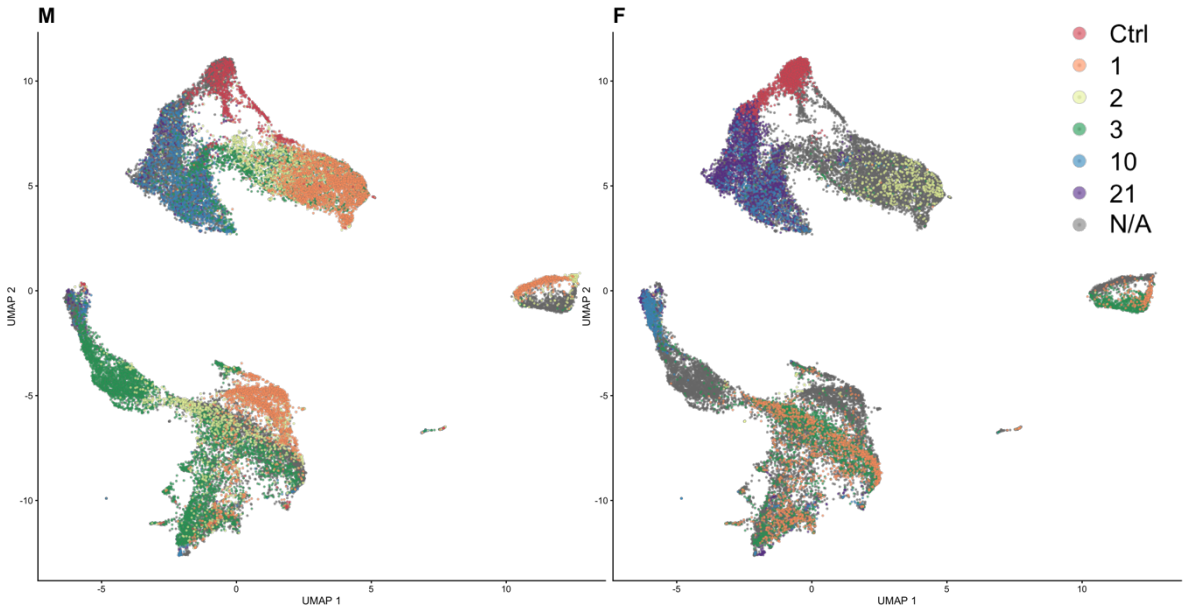


Figure 5.9 UMAPs of male (left) and female (right) myeloid cells. Each cell is coloured by its dpi. Male and female samples were not perfectly replicated across the dataset, so it's difficult to draw conclusions. However, there may be slight differences between the sexes.

5.3 Unsupervised Clustering

After establishing that the myeloid cell atlas exhibited the expected cell types and states based on previous findings, I sought to identify new transcriptional patterns involved in the SCI immune response by employing unsupervised analytical approaches. To this end, I performed unsupervised clustering on the myeloid cell atlas via Leiden community detection (Traag et al., 2019) by converting the *SingleCellExperiment* object to a *Monocle3 cell_data_set* and calling the *cluster_cells* function from the *Monocle3* (version 0.2.1) with the parameters *reduction_method* = UMAP and *k* = 14. Briefly, this approach first builds a k nearest neighbour graph with the number of nearest neighbours defined by the parameter *k* (Amezquita et al., 2020). Notably, the value of *k* regulates clustering resolution, with greater values resulting in more connectivity and ultimately broader clusters, while lower values have the opposite effect. Here, the value of *k* was empirically selected. After building the graph, *cluster_cells* calculates the weight of the connection between each cell using the Jaccard index, and then identifies highly connected cells via Leiden community detection. In this chapter, I chose to use the Leiden over the Louvain method applied in Chapters 3 and 4 because it is faster, more

effective at generating well-connected communities, and is the recommended approach when performing *Monocle3* trajectory analysis (Nayak & Hasija, 2021; Traag et al., 2019).

I visualised the clusters on a UMAP (Figure 5.10A) and investigated the correlation between dpi and cluster membership. I found highly statistically significant correlations for clusters 1-15 (Figure 5.10B) suggesting a level of biological relevance for these clustering results. Conversely, clusters 17-22 were only weakly correlated with dpi and were also the smallest clusters (< 165 cells). Cluster 22 (< 20 cells), was likely a technical artifact as it originated from a single sequencing flow cell. Cluster 15, while well correlated with the Ctrl, comprised cells from predominantly one collection date (6-Mar-20), had the lowest logcounts and highest mitochondrial gene percentage, and was marked by DEGs that were not typical of microglia (*Cacna1d*, *Mlxipl)* (Franzén et al., 2019). This cluster likely represented low quality cell libraries and was not considered for further analysis.

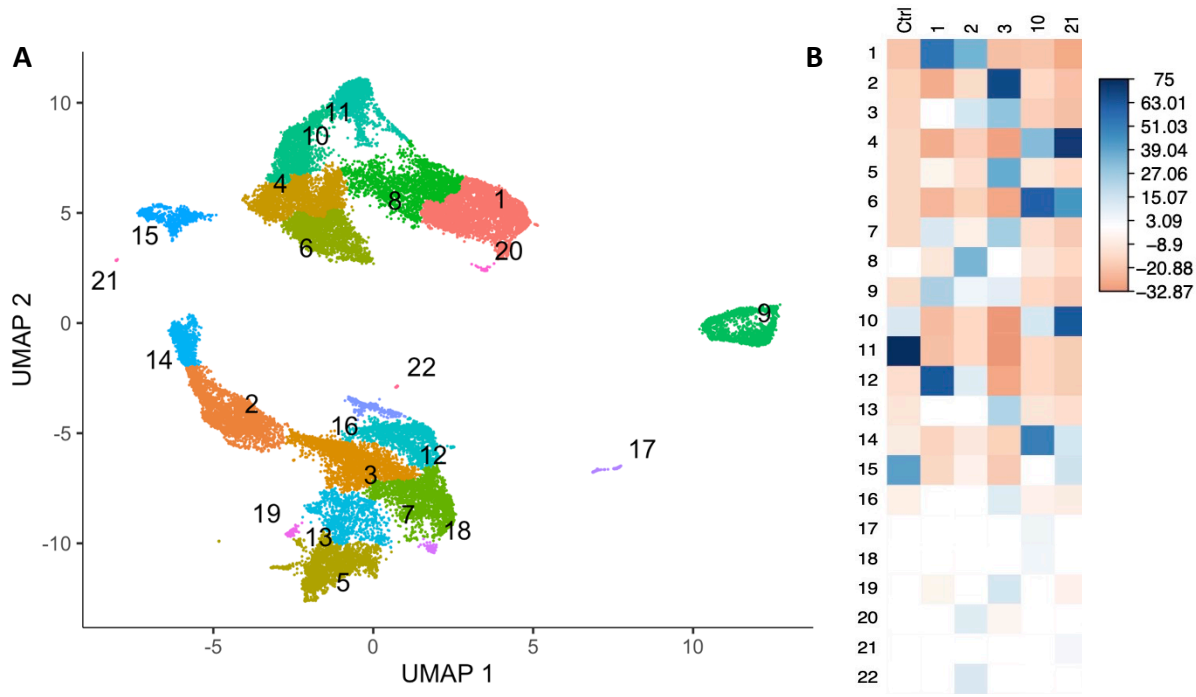


Figure 5.10 UMAP of the myeloid cell atlas. Each dot represents a cell and is coloured by cluster membership, as determined by Leiden community detection using *Monocle 3*. B) A correlation plot of the Pearson's chi-squared residuals between cluster membership (rows) and dpi (columns). Pearson's Chi-squared test; X-squared = 60992, df = 105, p-value < 2.2e-16. Blue signifies a positive standardised residual, while red signifies the opposite. White represents a non-significant contribution. Note that the cluster 11 – Ctrl standardised residual was limited to maintain contrast, its true value was 142. Certain clusters, such as cluster 11, are highly correlated with a single stage, while others, like cluster 4, are split across more than one stage. Clusters 17-22 are only weakly correlated with dpi.

Next, I looked at the relationship between cell cycling and cluster membership, expecting cells of similar cell types and cycling stages to cluster together. In fact, for many clusters (e.g., 1, 2,

6, 8, 11) there were strong correlations (Figure 5.11). Clusters 17-22 were again only weakly correlated, stemming from their small size, ultimately making them less reliably interpreted. The three smallest clusters, 20, 21 and 22 (< 55 cells), were not considered in the downstream analysis.

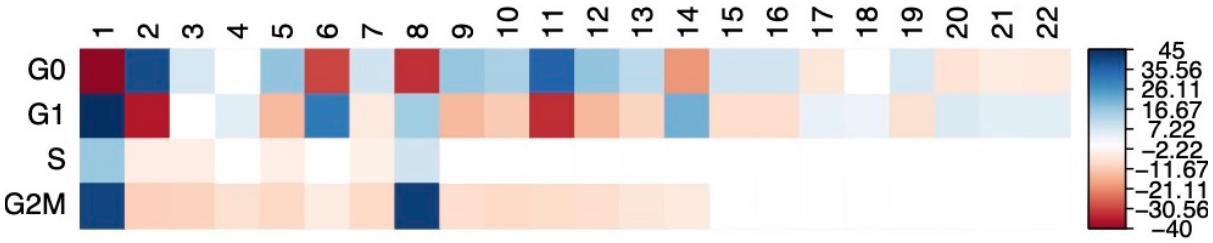


Figure 5.11 A correlation plot of the Pearson's chi-squared residuals between cell cycle (rows) and cluster membership (columns). Pearson's Chi-squared test; X-squared = 13837, df = 63, p-value < 2.2e-16. Blue signifies a positive standardised residual, while red signifies the opposite. White represents a non-significant contribution. Note that the cluster 1 – G0 and 1 – G1 standardised residual was limited to maintain contrast. True value were -74 and 57, respectively. In many cases, cell cycle and cluster membership are correlated.

5.4 Trajectory Analysis

After noting the correlations between cluster membership and dpi or cell cycle phase, I decided to perform unsupervised trajectory analysis to further investigate the transition between cell states within the myeloid cell atlas. To this end, I employed the *Monocle3* function *learn_graph* on the *cell_data_set* object with the parameter *close_loop* = false (Cao et al., 2019; Qiu et al., 2017; Trapnell et al., 2014). Briefly, this function uses reversed graph embedding to learn the changes in gene expression that could signify a transition between cell states (Cao et al., 2019; Qi et al., 2017). The resultant trajectory can be observed in Figure 5.12A. Given that I collected time-resolved samples, I was able to use this information to provide directionality to the trajectory. To visualise this, I called the *Monocle 3 order_cells* function, which aligns each cell along the learned trajectory, assigning the cell a value in *pseudotime*. The *pseudotime* can then be given directionality by manually selecting trajectory roots corresponding to cells from 1 dpi (Figure 5.12B). Despite only providing the algorithm with information about 1 dpi, the *pseudotime* and dpi were well aligned, creating a trajectory for the myeloid cell transcriptional profiles across the acute, subacute, and early chronic phases of SCI.

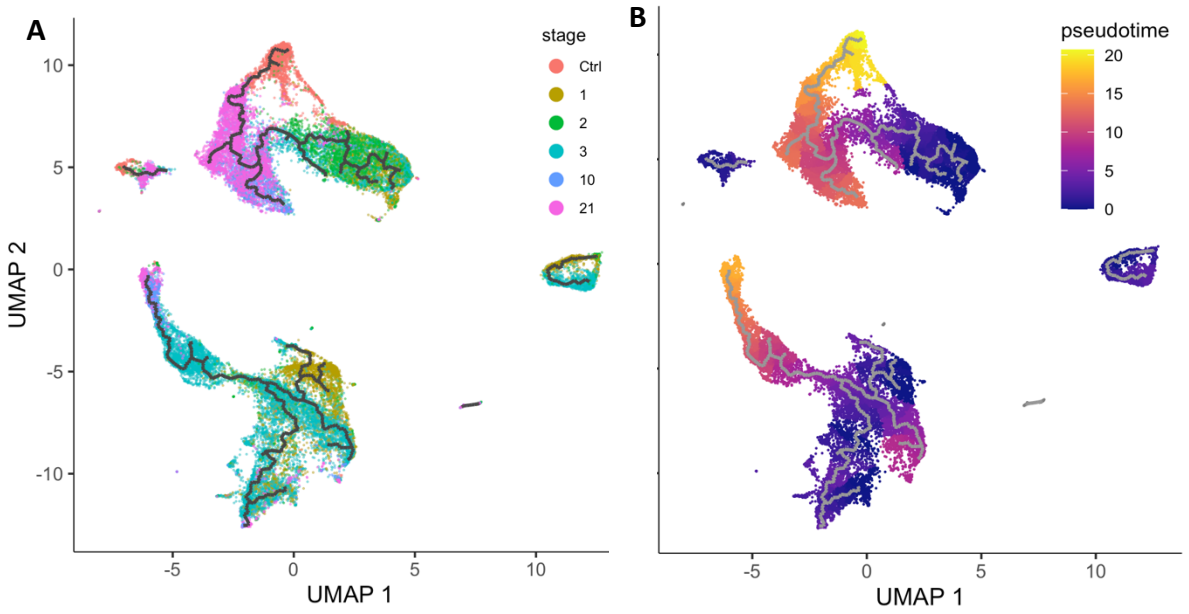


Figure 5.12 UMAP of the myeloid cell atlas coloured by dpi A) or *Monocle 3* inferred pseudotime B). The UMAPs are superimposed with the unsupervised reversed graph embedding-derived trajectory inferred via *Monocle 3*. Pseudotime and dpi are closely aligned.

5.5 Characterising Myeloid Cell Atlas Clusters

To ultimately infer the biological meaning of clusters in a scRNAseq dataset, one can calculate the DEGs that characterise each cluster and extrapolate possible functions using GO analysis and published mechanistic studies. This can then be harmonised with other atlas-derived data, such as the cell cycling annotations and trajectory analysis, to infer a map of myeloid cell transcriptional profiles across the SCI pathology. To begin, I provided the clusters with more intuitive names based on their majority cell-type and their position in the trajectory (Figure 5.13).

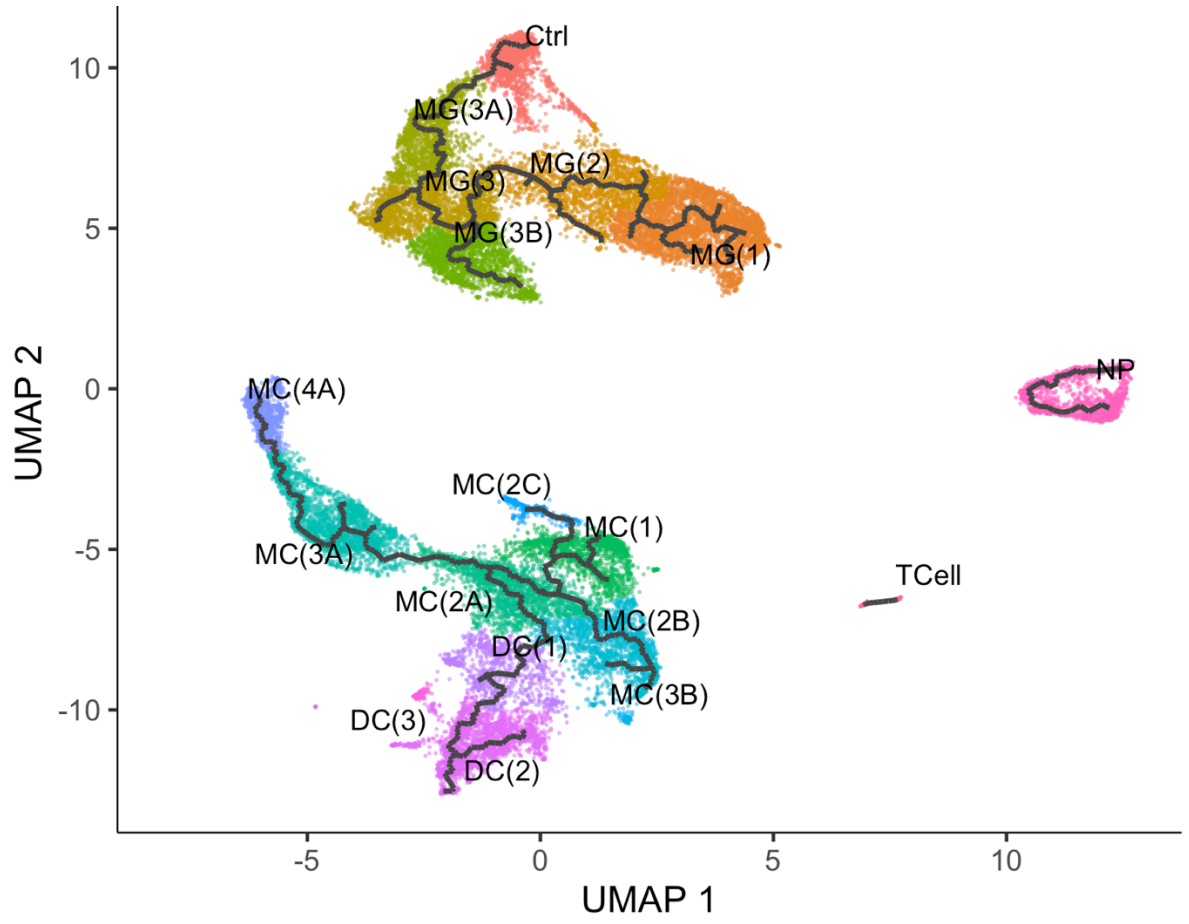


Figure 5.13 UMAP of the myeloid cell atlas coloured by the updated cluster names. The UMAP is superimposed with the unsupervised reversed graph embedding-derived trajectory inferred via *Monocle 3*. MG = microglia, MC = monocyte/macrophage, DC = dendritic cells, NP = neutrophils.

5.5.1 Calculating DEGs and Performing GO Analysis

I generated DEGs for each cluster by calling the *scran* function, *findMarkers*, with the parameters *group* = cluster membership, *direction* = upregulated, *pval.type* = some. The *findMarkers* function identifies DEGs by pairwise Welch t-Tests comparisons corrected for multiple testing via the Benjamini-Hochberg method and combines the comparisons into a ranked list of markers for each cluster (Figure 5.14). I then used these DEGs for GO enrichment analysis. To this end, I selected the top DEGs for each cluster, defined as FDR < 0.0001, FC > 1.5, ranging from 77-185 genes). Then, I submitted them to the PANTHER Classification System overrepresentation test (Mi et al., 2013) (release 2021-07-02) following the instructions on the Gene Enrichment analysis page (http://geneontology.org/docs/go-enrichment-analysis/), including the optional use of a custom reference list, which was the detected genes from the custom mm10 reference list. P-values were computed using the default setting of Fisher's Exact Test and were corrected for multiple testing using the

Benjamini-Hochberg method. I examined only biological processes that were over-represented.

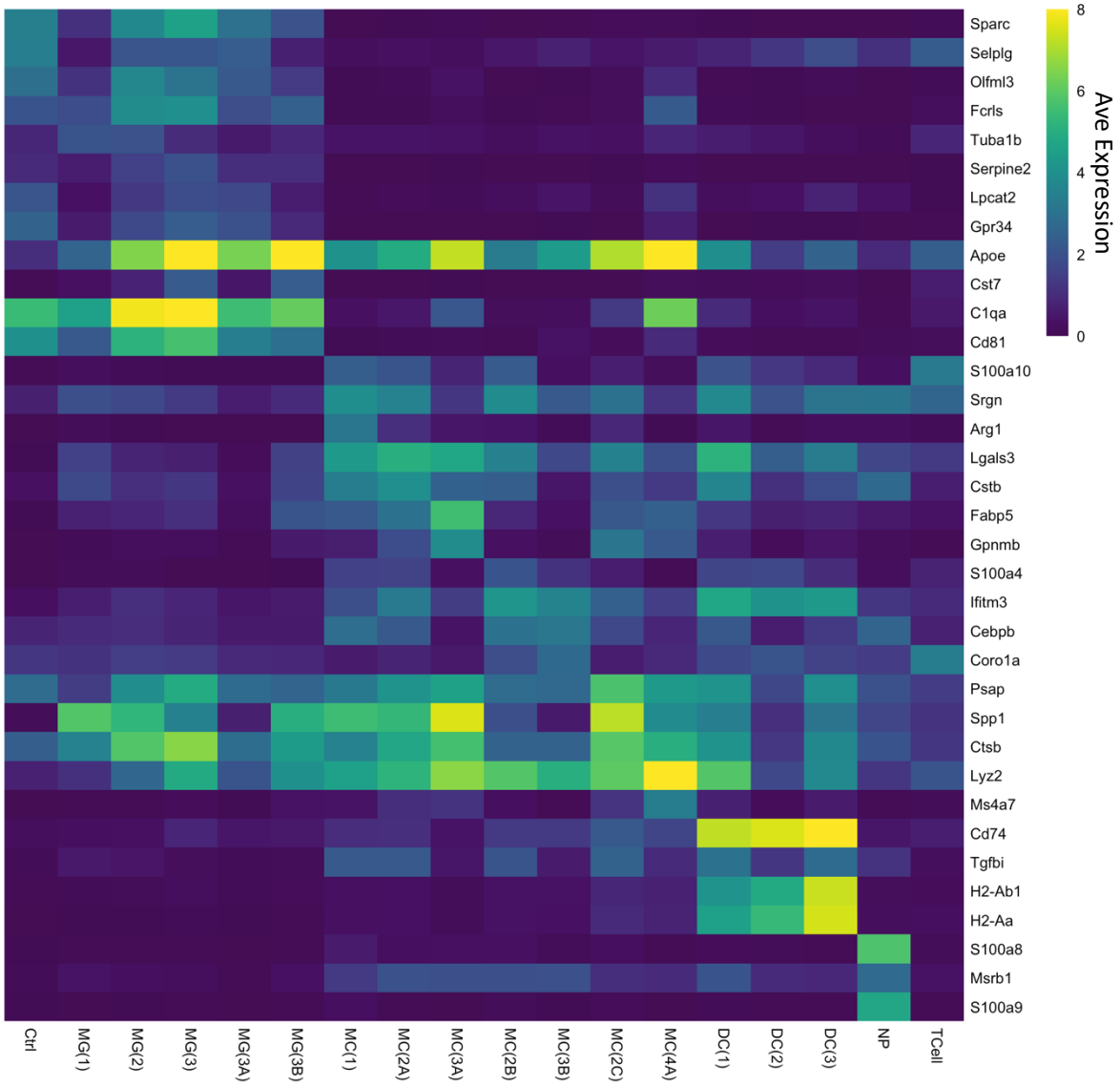


Figure 5.14 Heatmap of average expression of the top 3 DEGs per cluster. Note that some top DEGs overlapped, and the 6 cells coloured bright yellow (max) were scaled down to maintain contrast; their true values ranged from 9-13 logcounts.

5.5.2 Characterising Microglia Clusters

At 1 dpi, the majority of microglia were in a highly activated MG(1) state (Figure 5.16A) characterized by GO terms for proliferation, as expected from the cell cycle analysis (Figure 5.16B) and recent reports (Bellver-Landete et al., 2019) (Figure 5.16C). As previously demonstrated, these cells downregulated homeostatic microglia markers (Figure 5.15A) and began to adopt DAM1-like phenotypes (Figure 5.5; Figure 5.15B)

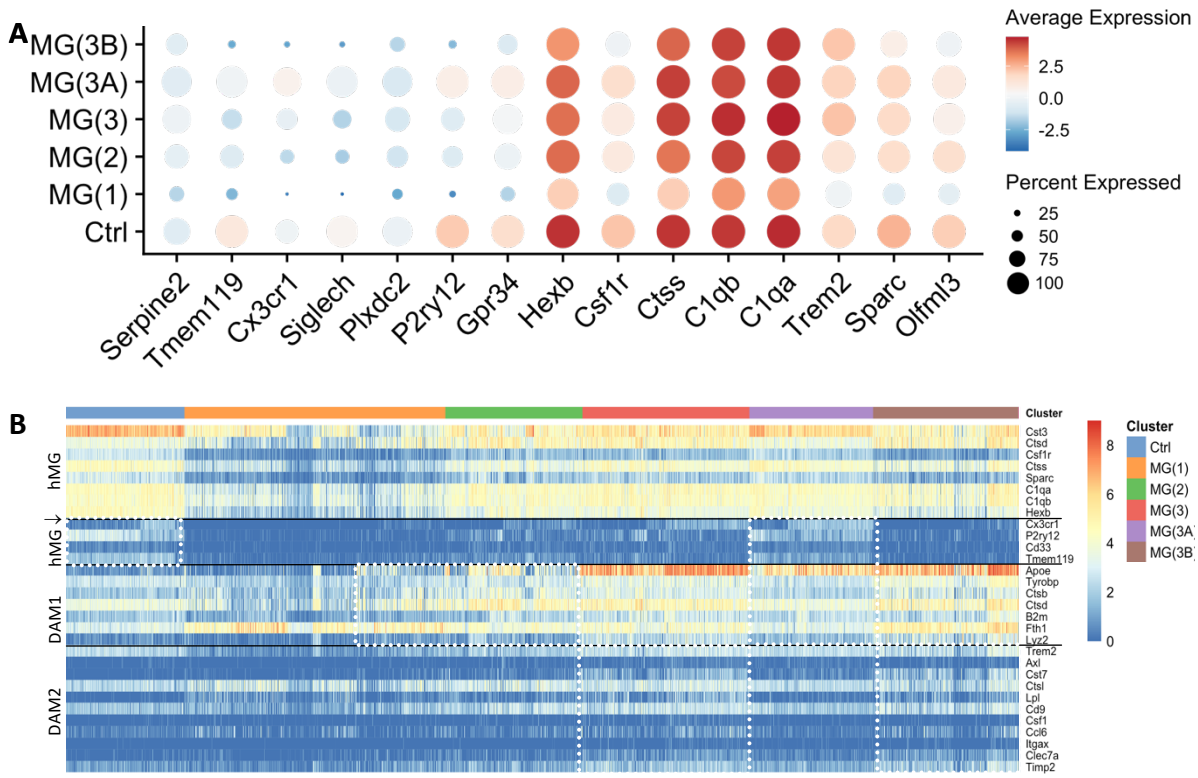


Figure 5.15 A) Dot plot showing canonical microglia marker expression across the microglia clusters. Homeostatic genes are downregulated as early as cluster MG(1). However, by MG(3A), many homeostatic microglia gene levels closely resembles the Ctrl. B) A heatmap of homeostatic and DAM-associated genes in microglia across the different clusters. Profiles are separated by solid black line, hMG = homeostatic microglia, hMG↓= homeostatic microglia that are downregulated in the DAM1 signature. Each row is a gene, and each column is a cell. Dotted white boxes highlight microglia that fit the designated profile.

MG(1) was also highly enriched for terms regarding the production of and response to cytokines, as well as signs of aerobic glycolysis, which has been shown to support cytokine production and is a hallmark of pro-inflammatory macrophages (Meiser et al., 2016). MG(1) was also enriched for GO terms for the recruitment of blood-borne leukocytes and migration as described (Milich et al., 2021), positive regulation of astrocyte activation (Donnelly & Popovich, 2008), and lipid storage and metabolism (Figure 5.16C). MG(1) notably lacked colony-stimulating factor 1 (*Csf1r*), a gene required for microglia proliferation and survival during homeostasis (Elmore et al., 2014) and differentially upregulated epidermal fatty acid-binding protein 5 (*Fabp5*), a target of both the NF-κB and PPAR-γ pathways that is involved in lipid transport/localization (Kaczocha et al., 2014; Senga et al., 2018). The upregulation of *Fabp5* appeared to occur at the root of the MG(1) trajectory that passed through the highly proliferating population (Figure 5.7B). This is in line with reports that *Fabp5* is both required for proliferation and, under polyunsaturated fatty acid-high conditions, can bind HIF-1α to facilitate proliferation (Levi et al., 2013; Seo et al., 2020; Zhao et al., 2017).

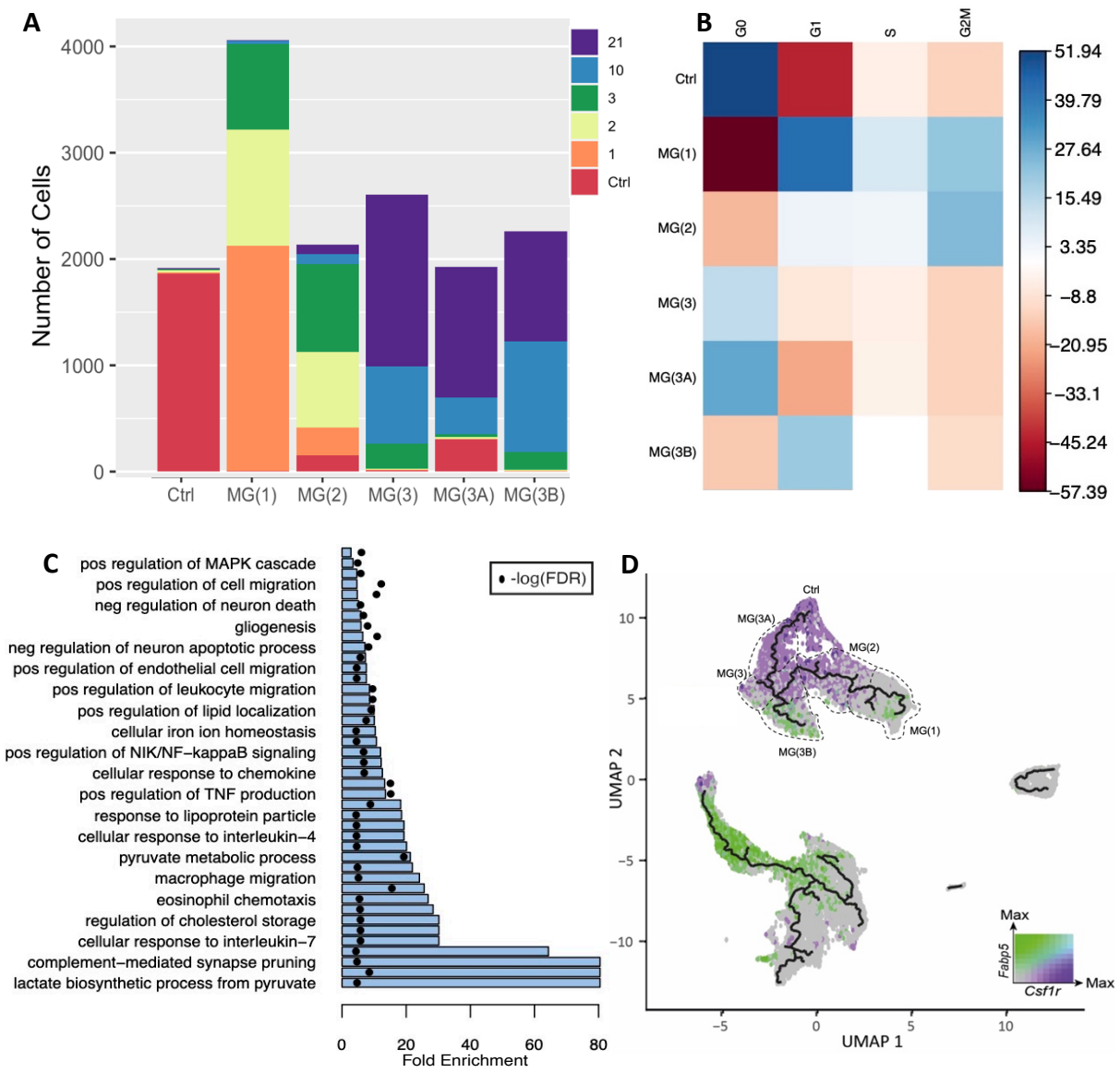


Figure 5.16 A) Bar chart of the microglia clusters depicting the number of cells per cluster, coloured by their dpi. B) A correlation plot of the Pearson's chi-squared residuals between microglia cluster membership (rows) and cycle (columns). Pearson's Chi-squared test X-squared = 7417.2, df = 18, p-value < 2.2e-16. C) The fold enrichment and FDR of select GO terms for cluster MG(1). Pos = positive; Neg=negative. D) UMAP coloured by the scaled logcounts of *Fabp5* and *Csf1r*. Note the lack of blue, demonstrating that these two genes are not co-expressed and MG(1) *Fabp5* is expressed at only one of the MG(1) trajectory roots. The UMAP is superimposed with the trajectory inferred via *Monocle 3*. Microglia clusters are denoted by the dotted lines.

By 2 dpi, about half of microglia remained in the MG(1) stage, while the others adopted a MG(2) phenotype (Figure 5.16A), which closely resembled the DAM1 phenotype (Figure 5.5; Figure 5.15B) (Keren-Shaul et al., 2017). MG(2) cells were more likely to be in G2M phase of proliferation (Figure 5.16B). These cells also downregulated *Fabp5* and upregulated *Csf1r* (Figure 5.16D). According to GO term analysis, this population continued to produce cytokines and perform aerobic glycolysis (Figure 5.17A), migrate, activate astrocytes, promote angiogenesis, recruit blood borne leukocytes, and metabolise lipids. Unlike MG(1), MG(2) also

showed signs of phagocytosis and began to upregulate *Apoe* and *Trem2* (Figure 5.17B,C), adding further evidence for phagocytic functions (Atagi et al., 2015).

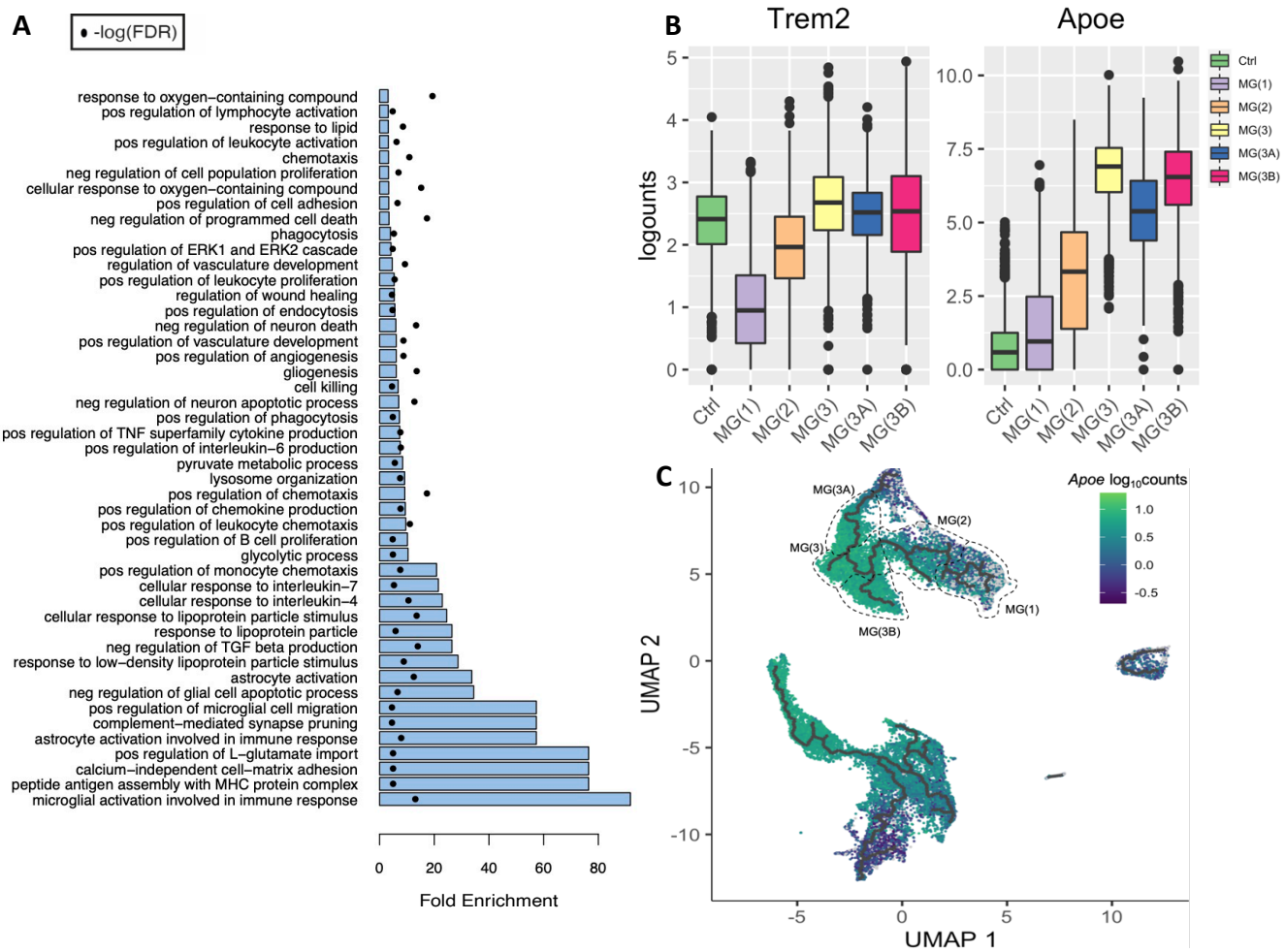


Figure 5.17 A) The fold enrichment and FDR of select GO terms for cluster MG(2). Pos = positive; Neg=negative. B) Boxplots of *Trem2* and *Apoe* logcounts by microglia cluster. MG(2) begins to restore homeostatic *Trem2* levels and upregulates *Apoe*, suggesting an increase in lipid metabolism, possibly due to phagocytosed myelin debris. C) UMAP coloured by the log_{10}counts of *Apoe.* The UMAP is superimposed with the trajectory inferred via *Monocle 3*. Microglia clusters are denoted by the dotted lines. *Apoe* is quite differentially expressed across the myeloid cell atlas, with the highest microglia expression in the MG(3) clusters.

To further verify that MG(1) and (2) cells were proliferating, I investigated the expression of *P2ry12, Msr1,* and *Cdk1.* In a recent scRNAseq study of CNS cells after SCI, *P2ry12* was reported to be downregulated, while *Msr1* and *Cdk1* were upregulated in cycling microglia (Milich et al., 2021). I observed this pattern in MG(1) (Figure 5.18A) and, to a lesser extent, in MG(2) (Figure 5.18B), which began to upregulate *P2ry12.* Compared to other clusters, however, *Msr1* and *Cdk1* expression was increased (Figure 5.18C). I also confirmed the presence of previously described proliferating microglia profile and found that MG(1) and (2) upregulated the expected genes (Figure 5.18D) (Hammond et al., 2019; Milich et al., 2021).

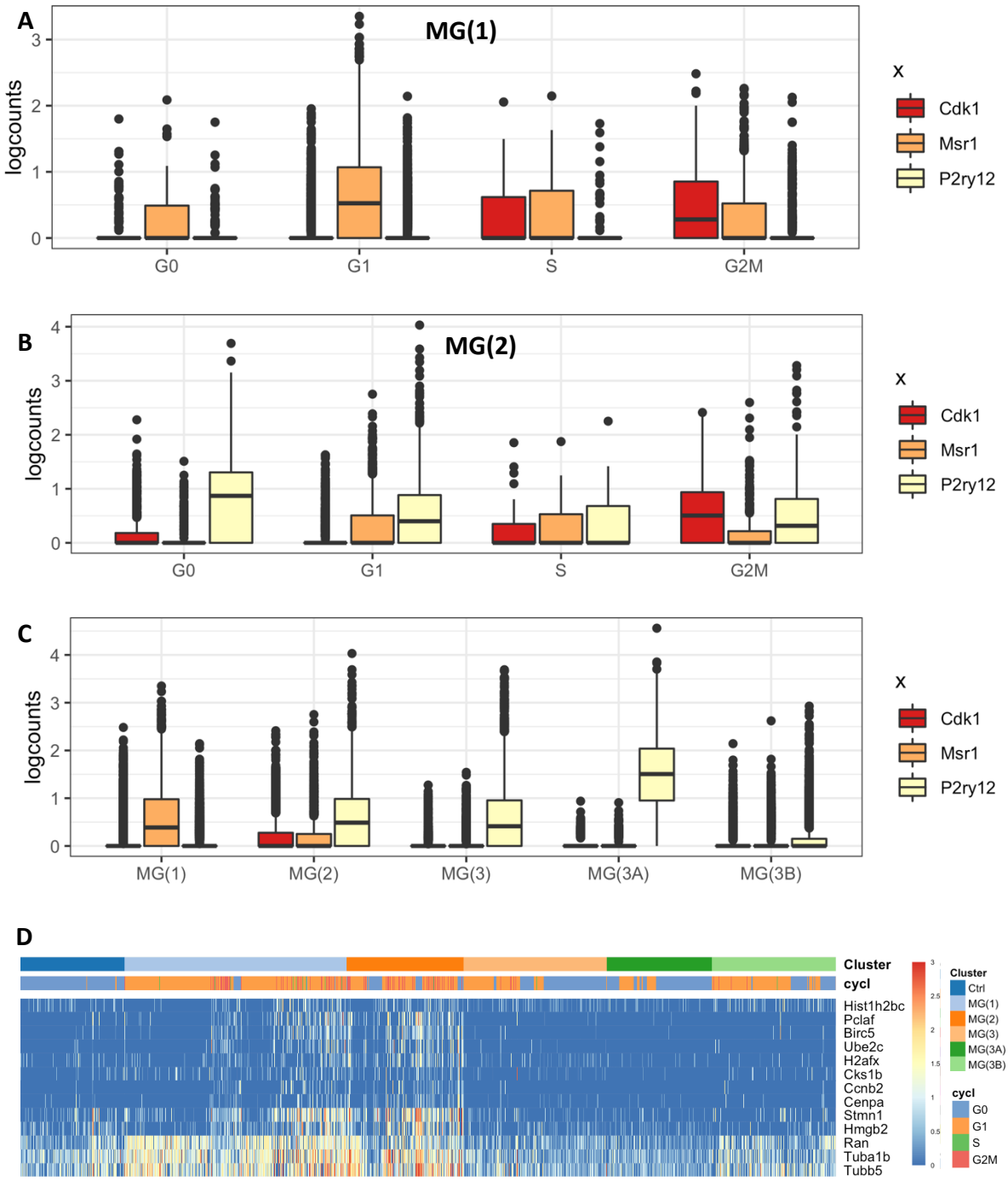


Figure 5.18 Boxplots of *Cdk1, Msr1,* and *P2ry12* expression in A) MG(1) or B) MG(2) across cell cycle phases or C) across all activated microglia clusters. *Cdk1* and *Msr1* expression loosely mark the two acute cycling clusters. D) Heatmap of microglia from the myeloid cell atlas and their expression of proliferating microglia-associated genes. Each row is a gene, and each column is a cell. Clusters MG(1) and (2) strongly upregulate proliferation profiles.

3 dpi, microglia remained largely split between the proliferating, proinflammatory clusters, MG(1) and MG(2), with ~11% clustering in MG(3) (Figure 5.16A). MG(3) adopted a more DAM2-like phenotype (Figure 5.15B) but still showed GO terms for blood-borne leukocyte recruitment as well as phagocytosis and lipid processing, supported by high expression levels of lipid processing genes *Trem2, Apoe,* and *Fabp5* (Figure 5.17B, Figure 5.19A,B). MG(3)

continued to activate astrocytes and began to upregulate *Igf1*, which has been shown to promote astrogliosis and ultimately scar formation in SCI (Figure 5.19A,B) (Bellver-Landete et al., 2019). Unlike MG(1/2), MG(3) was no longer proliferating (Figure 5.7B; Figure 5.16B; Figure 5.19A), and had GO terms for cytotoxicity, ROS-producing reverse electron transport (Peruzzotti-Jametti et al., 2021), and the production of IL-6, a cytokine known to sustain chronic inflammation, possibly by shifting monocyte differentiation from a dendritic cell to macrophage fate (Chomarat et al., 2000; Hunter & Jones, 2015). Conversely, astrocyte-targeted IL-6 has been shown to promote corralling of pro-inflammatory infiltrating cells, which together with the observed upregulation of scar-formation promoting, *Igf1* (Figure 5.19B), suggests a role for scar formation (Bellver-Landete et al., 2019; Penkowa et al., 2003). Notably, this cluster also presented several contradictory GO terms, such as both the positive and negative regulation of cytokine production (Figure 5.16D). These discrepancies might be explained by the heterogeneity within MG(3), depicted by the branching trajectory that passes through this cluster, including the differential expression of *Fabp5* and *Csf1r* within the cluster (Figure 5.16A).

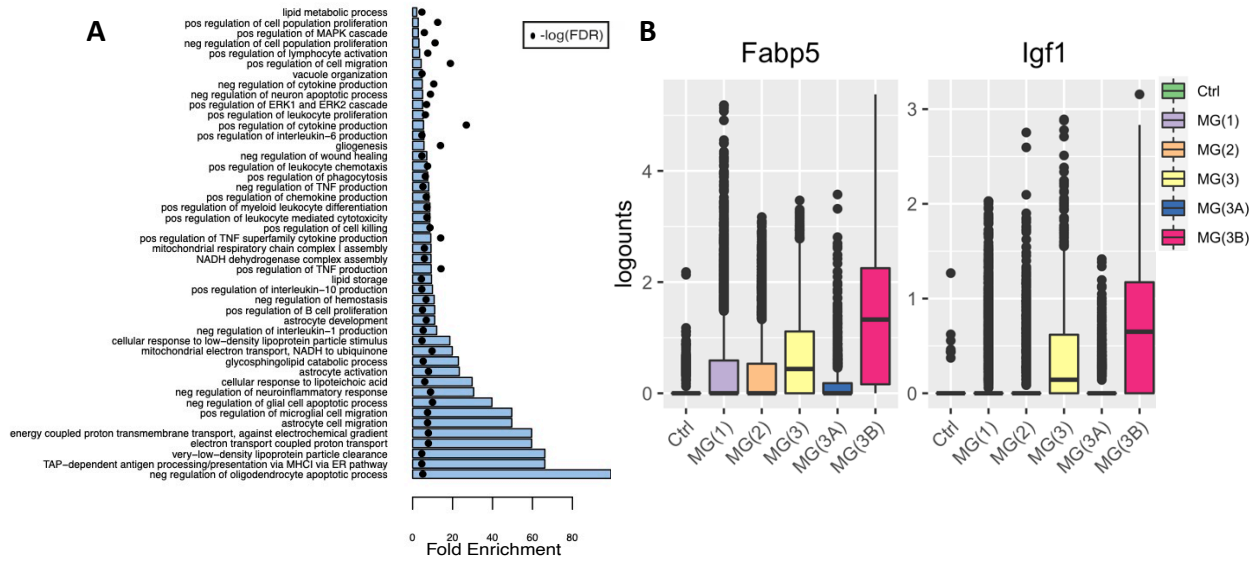


Figure 5.19 A) The fold enrichment and FDR of select GO terms for cluster MG(3). Pos = positive; Neg=negative. B) Boxplots of *Fabp5* and *Igf1* logcounts by microglia cluster. MG(3) upregulates both genes, suggesting an increase in lipid metabolism and the beginning of scar formation.

At 10 and 21 dpi, microglia were split between MG(3), MG(3A), and MG(3B) (Figure 5.16A). Of all the microglia clusters, MG(3A) most closely resembled the homeostatic Ctrl cluster in regard to its homeostatic gene expression (Figure 5.15A), cell cycling phase (Figure 5.16B),

trajectory and pseudotime (Figure 5.12), and GO Terms (Figure 5.20A,B). GO terms for both the Ctrl and MG(3A) included antigen presentation, leukocyte homeostasis, phagocytosis, synapse pruning, cell adhesion, and motility. A major difference between the Ctrl and MG(3A) clusters, was the upregulation of *Apoe* (Figure 5.17B,C) and other DAM1-genes such as *Tyrobp* and *Lyz2* (Figure 5.15B). Compared to MG(3) and MG(3B), the DAM2-like expression pattern was noticeably absent in MG(3A).

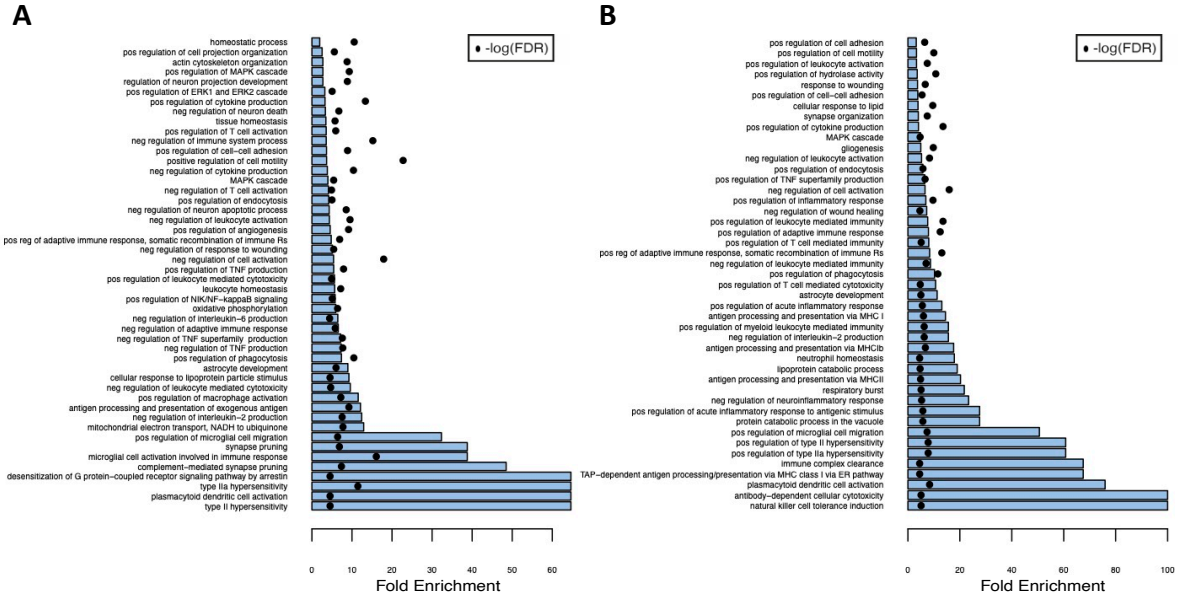
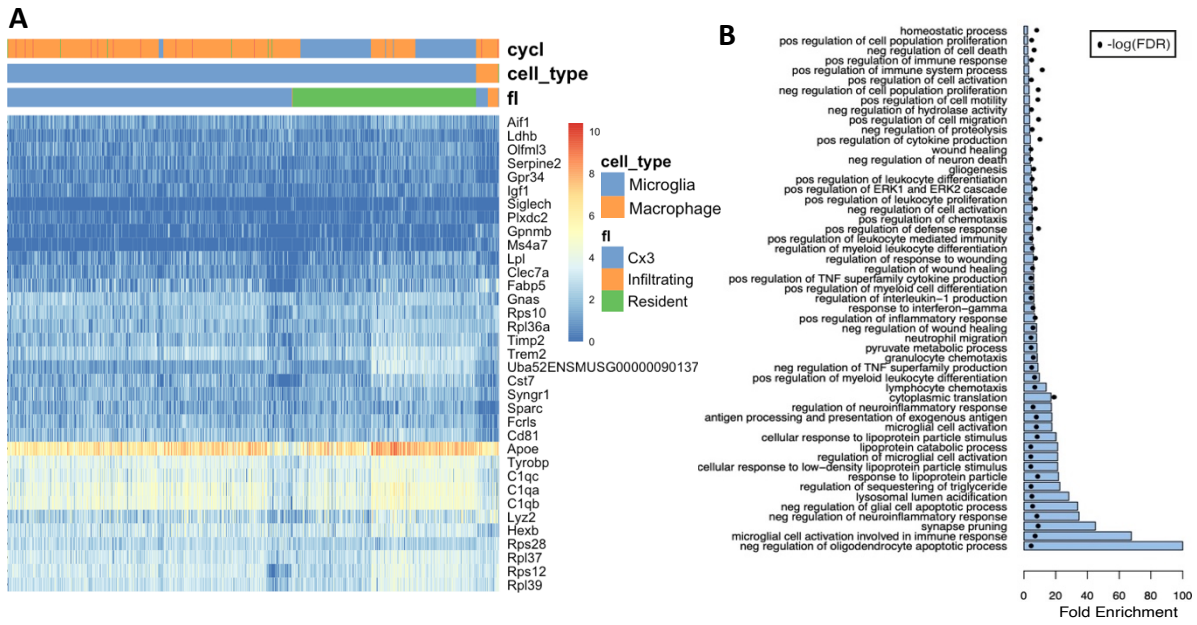


Figure 5.20 A) The fold enrichment and FDR of select GO terms for the Ctrl cluster. B) The fold enrichment and FDR of select GO terms for the MG(3A) cluster. Pos = positive; Neg=negative; Reg = regulation; Rs = receptor. The Ctrl and MG(3A) clusters share many terms.

MG(3B) comprised ~8% 3 dpi cells but was predominantly populated by cells from 10 and 21 dpi. Notably, not all of these cells were microglia. Approximately 5% were macrophages and could be distinguished by their fate mapping label (for the *CreMato* samples), their differential expression of microglia markers, particularly *Siglech, Serpine2,* and *Sparc*, and the upregulation of *Ms4a7* (Figure 5.21A) in line with previous reports (Jordão et al., 2019; Konishi et al., 2017). Like MG(3), MG(3B) adopted a more DAM2-like phenotype (Figure 5.15B) and downregulated homeostatic microglia genes, even below MG(3) levels (Figure 5.15A). Unlike MG(3A), MG(3B) showed some signs of proliferation (Figure 5.7B; Figure 5.16B), but not as convincingly as the acute phase clusters (Figure 5.18D) and the MG(3B) cluster also had the highest expression of scar-formation promoting, *Igf1* (Figure 5.19B) (Bellver-Landete et al., 2019). In terms of GO analysis, cluster MG(3B), like MG(3) and MG(3A), showed signs of migration, but lacked antigen presentation, metabolism-related genes, and phagocytosis (Figure 5.21B). GO terms for processing phagocytic cargo, such as lysosomal acidification and

triglyceride metabolism, were enriched (Chandak et al., 2010; Marschallinger et al., 2020). Genes for lipid processing were also upregulated including *Apoe* (Figure 5.17B,C) and *Fabp5* (Figure 5.19B; Figure 5.16D).

I wondered whether the lipid-processing GO terms reflected a foam-cell like phenotype, and investigated the expression of peripheral macrophage foam cell (Willemsen & de Winther, 2020; Zhu et al., 2017) and lipid-droplet accumulating microglia (LDAM) markers (Marschallinger et al., 2020). Interestingly, MG(3B) upregulated many foam-cell markers but did not match the LDAM profile (Figure 5.21C). Macrophage *Msr1* expression was recently demonstrated to promote foamy macrophage formation and neuronal apoptosis through myelin-mediated NF-κB signalling in weight-drop mouse models of T10 contusive SCI (Kong et al., 2020). Interestingly, *Msr1* was upregulated in 1 dpi cells from MG(1) and (2), suggesting that the foam cell-like state of MG(3B) is initiated very early on in the microglia trajectory (Figure 5.21D, E).



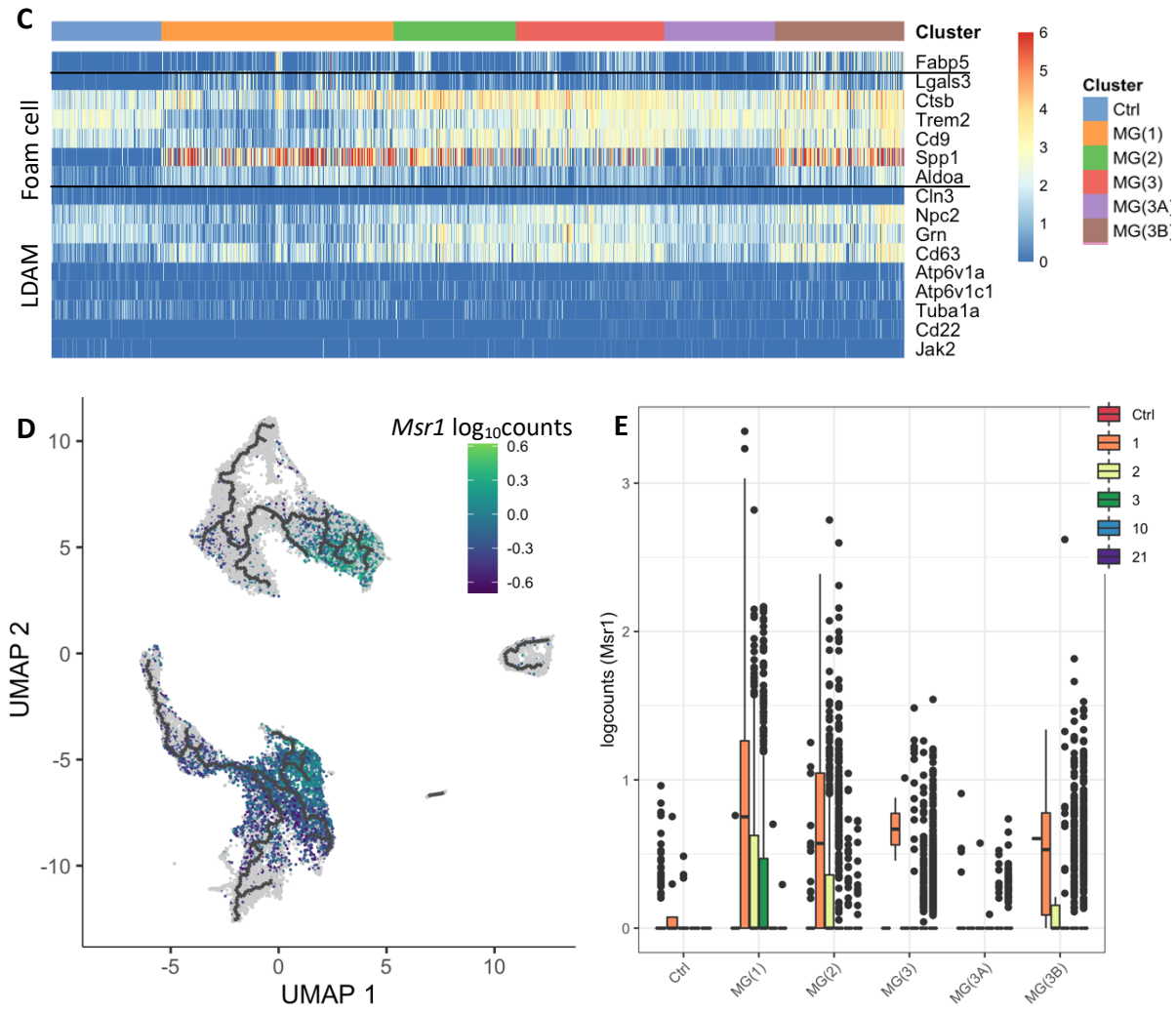


Figure 5.21 A) Heatmap of cluster MG(3B) depicting the logcounts for the top 30 DEGs genes plus *Siglech*, *Gpnmb*, *Ms4a7*, *Fabp5*, and *Lyz2*. Cells (columns) are labelled by their fate mapping label. Macrophages are only subtly distinguished from the microglia. B) The fold enrichment and FDR of select GO terms for the MG(3B) cluster. Pos = positive; Neg=negative. Phagocytosis is absent from the list, while GO terms for processing phagocytic cargo are enriched. C) Heatmap of *Fabp5,* foam cell-associated, and LDAM gene expression in the microglia clusters. MG(3B) upregulates *Fabp5* and foam cell-associated genes. D) UMAP coloured by the log_{10}counts of *Msr1.* The UMAP is superimposed with the trajectory inferred via *Monocle 3*. E) Boxplots depicting *Msr1* expression in the MG clusters across each dpi. *Msr1* expression is highest at 1 dpi.

As observed in cluster MG(1), *Fabp5* and *Csf1r* expression in cluster MG(3B) cells was mutually exclusive. We investigated this expression pattern at the protein level through quadruple IF of FABP5 and CSF1R in 21 dpi *CreMato* tissue, as described (4.2.3). We found that $CSF1R^-$ /$FABP5^+$ microglia and macrophages, likely corresponding to MG(3B) (Figure 5.16A,D), dominated the lesion core, while the majority of $CSF1R^+$ microglia, possibly MG(3A), remained perilesional (Figure 5.22).

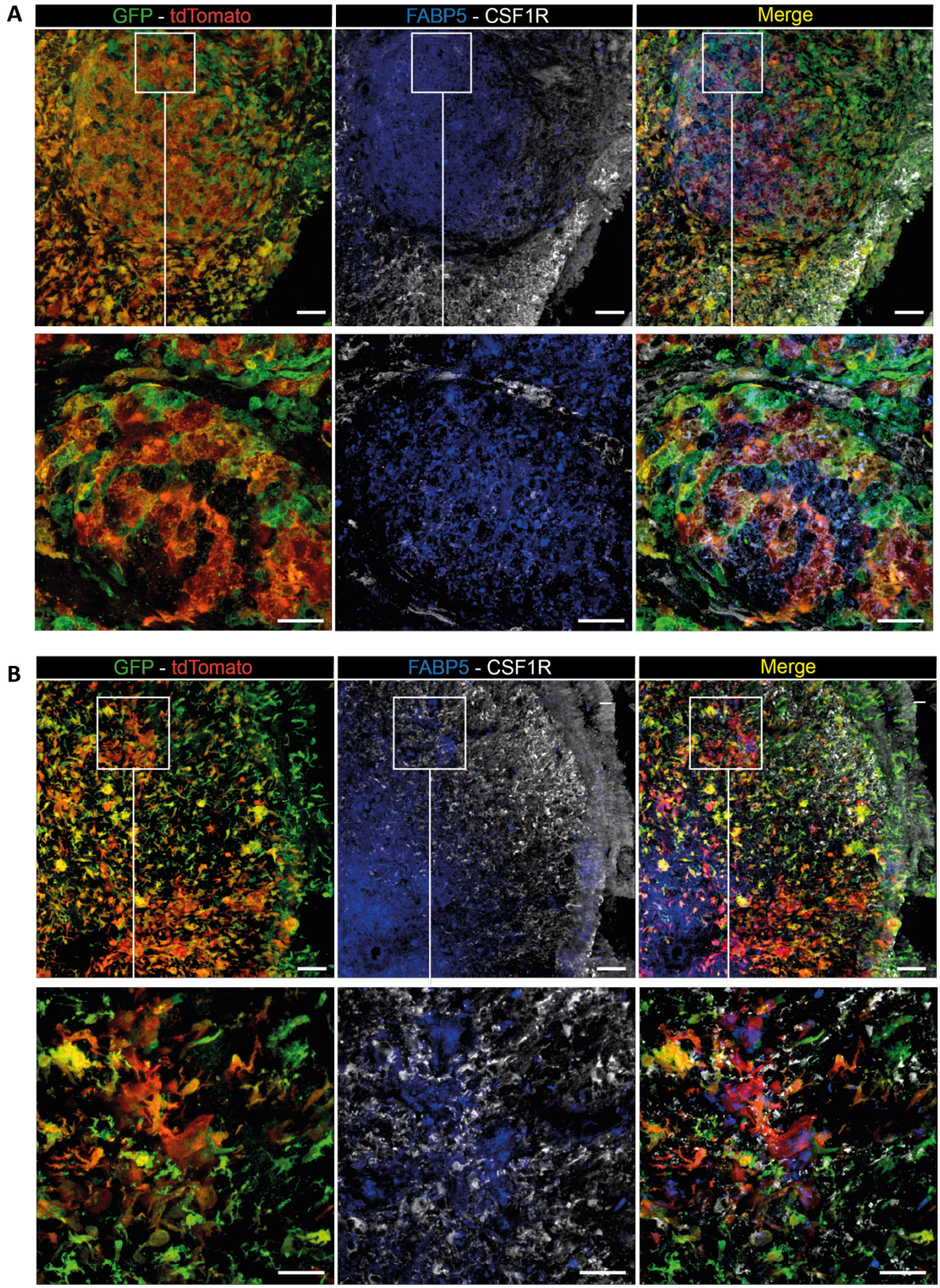


Figure 5.22 IF of *CreMato* tissue at 21 dpi stained for FABP5 and CSF1R at the lesion epicentre (A) and perilesional (B). Scale bars: Top panel 60 µm; Bottom panel 20 µm. FABP5 and CSF1R are mutually exclusive. FABP5 expression dominates the lesion core, while CSF1R is more prominent in the perilesional area. Images were acquired by Luca.

To further investigate the spatial dynamics of the microglia clusters across the SCI pathology, Veronica and I prepared SCI tissue and Katherine, Stefano, and I planned an smFISH experiment as described (4.2.2). Katherine performed the imaging and quantified spots in each region of interest (ROI). I performed the analysis on this data, the code for which can be found here (https://github.com/regan-hamel/PhD_Dissertation). To specifically identify the clusters *in situ*, we probed three genes, *Apoe*, *Trem2,* and *Fabp5*, plus the tdTomato and YFP fate mapping labels, ultimately defining each cluster according to Figure 5.23A. The average number of smFISH spots detected per cluster is shown in Figure 5.23B, and in general, closely reflects the scRNAseq data. MG(1) and (2) were difficult to match, likely since these clusters were not populous at the smFISH timepoints (3, 10, and 21 dpi).

Katherine segmented the tissue at the lesion epicentre into four concentric ROIs, centred on the lesion core, plus one ROI in the NASC (Figure 5.23C), allowing us to investigate the spatial dynamics of the clusters over time (Figure 5.23D). By 3 dpi, *Apoe*$^{-}$*/Fabp5*$^{-}$*/Trem2*$^{+}$ cells with Ctrl profiles were decreased at the lesion core but by 21 dpi, began to repopulate the core in greater numbers than the uninjured tissue, suggesting that a subset of proliferating microglia return to a homeostatic-like state in the chronic phase. In the scRNAseq data, I observed a similar trend but in the *Apoe*$^{+}$*/Fabp5*$^{-}$*/Trem2*$^{+}$ MG(3A) cluster. This discrepancy could reflect an isolation-induced upregulation of *Apoe*, as described (Milich et al., 2021). Activated microglia, were present across all the ROIs from 3 dpi. As previously mentioned, MG(1) and (2) were largely absent, especially near the lesion core. MG(3) was present at 3 and 10 dpi throughout the sampled tissue except at the core. MG(3B) accounted for large proportions of the cells across the tissue at 3 dpi, which is surprising given the relatively small proportion of 3 dpi MG(3B) cells in the scRNAseq data. However, this could be due to an isolation bias, which will be touched in the *Discussion* and further discussed in *6.1*. MG(3B) did persist to 21 dpi in the smFISH data. MG(3A) appeared to be consistently enriched across the tissue except in the 2nd ROI, suggesting they might not be a structural component of the microglial scar (Bellver-Landete et al., 2019).

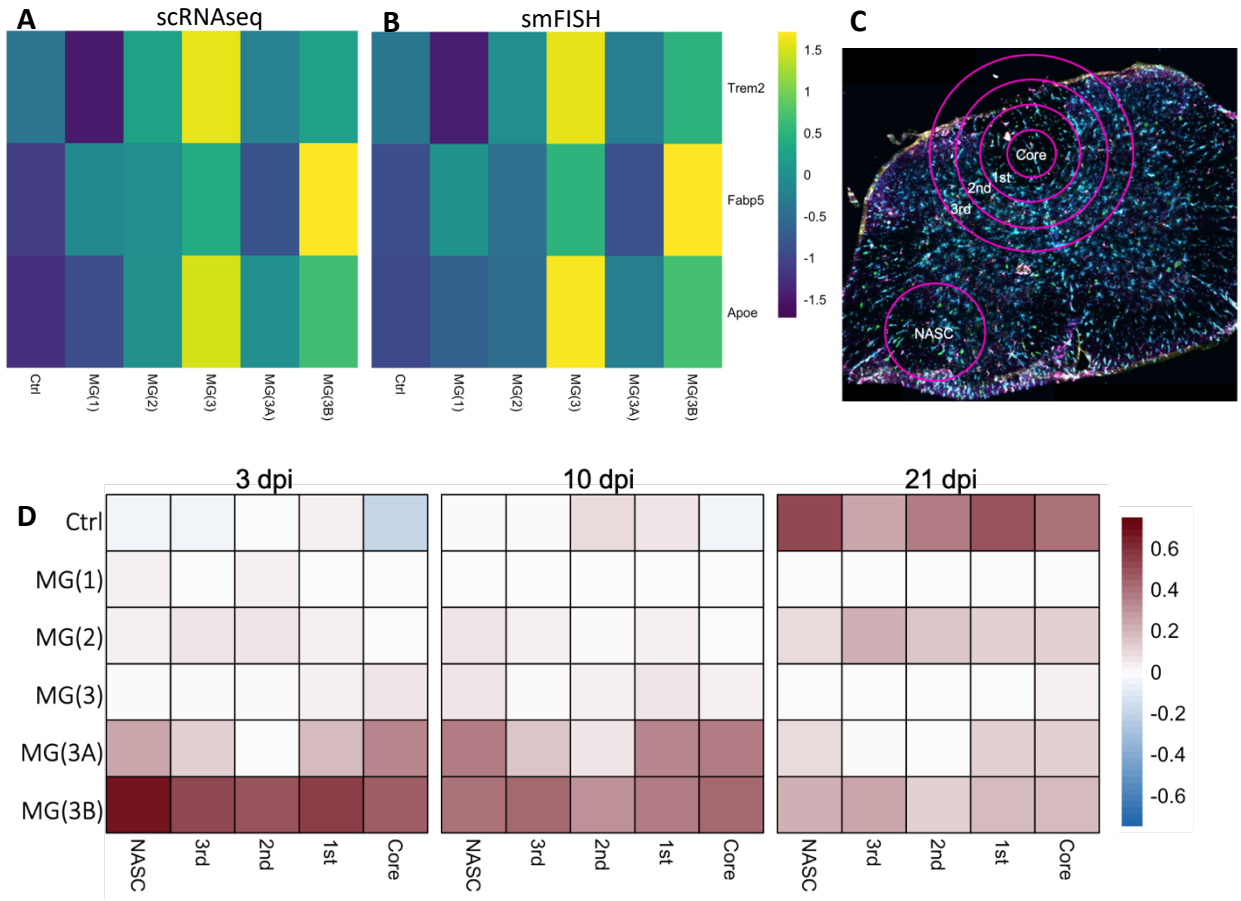
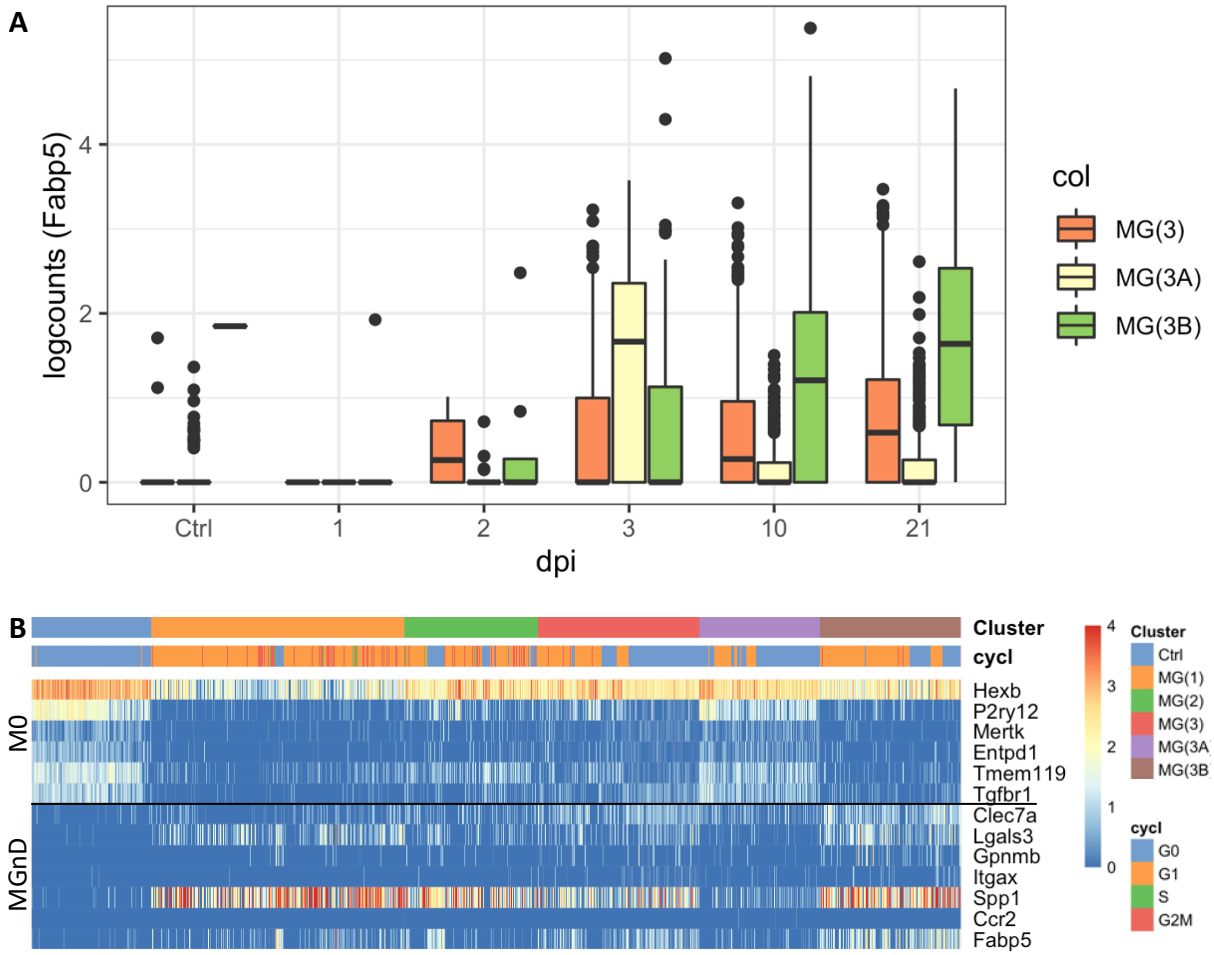


Figure 5.23 A) A heatmap of the scRNAseq data showing the average normalised expression of each gene of interest for each microglia cluster, scaled across each gene (row). B) A heatmap of the smFISH data depicting the average number of spots detected per cluster. The cluster definitions used for smFISH are well matched, with slight differences in cluster MG(1) and (2). These clusters were not well populated in the timepoints used for smFISH. C) The ROIs for the smFISH cluster quantification. This image was taken by Katherine. D) The smFISH-based quantification of microglia clusters by ROI at several dpi, presented as normalised residuals (condition – control/the number of microglia in the ROI).

The smFISH data, alongside the IF, support the *in silico* evidence for a dynamic *Fabp5* expression across the SCI pathology, particularly in the MG(3) trajectory. I investigated the expression level of *Fabp5* over time within the MG(3), (3A), and (3B) clusters (Figure 5.24A) and found that *Fabp5* increased over time for MG(3) and (3B). In (3A), which adopted a homeostatic-like phenotype post-SCI, only the 28 cells from 3 dpi expressed *Fabp5.* I wondered whether *Fabp5* expression might be negatively associated with the upregulation of microglial neuroprotective genes (M0), as defined by Krasemann et al., 2017. To address this, I first visualised the expression of neuroprotective genes across the clusters (Figure 5.24B) and observed an inverse pattern to the previously observed DAM2 gene expression (Figure 5.15B); the Ctrl and cluster MG(3A) upregulated neuroprotective genes, unlike the DAM2-like MG(3) and (3B) clusters. I also investigated MGnD defined by (Krasemann et al., 2017), which included *Fabp5,* and found that while both MG(3) and (3B) upregulated many

DAM2 genes, the latter was a better match with MGnD, suggesting a potential neurodegenerative role for this cluster. Next, I binned cells by *Fabp5* expression levels and compared the median expression of neuroprotective genes in each group. I found that *Fabp5*$^{hi/med}$ cells expressed lower levels of neuroprotective genes than *Fabp5*^{lo} cells (Figure 5.24C). Cumulatively, this data motivates the hypothesis that *Fabp5* upregulation may encourage microglia to follow the MG(3B) trajectory, rather than progressing to the homeostatic-like, neuroprotective MG(3A) state, a question which will be further deliberated in the *Discussion*.



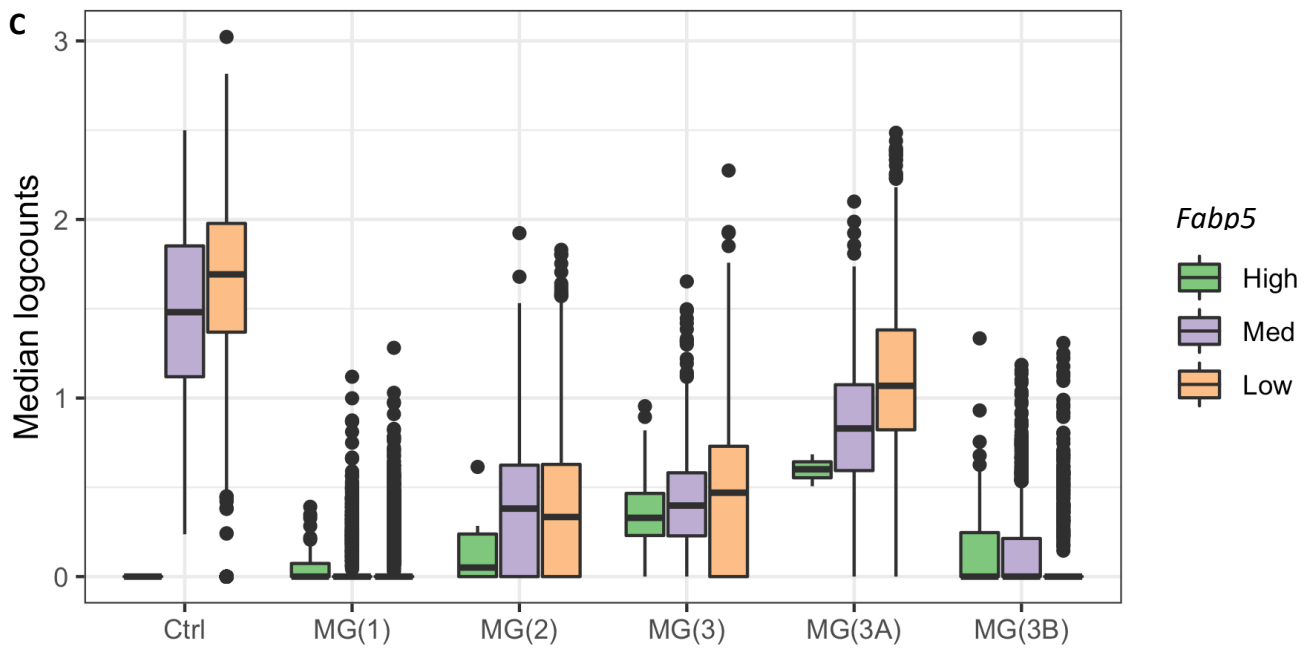


Figure 5.24 A) Boxplot of *Fabp5* expression in the MG(3) clusters over time. *Fabp5* expression increases over time except in cluster MG(3A). B) Heatmap of the microglia clusters depicting the logcounts of neuroprotective vs neurodegenerative genes. Cells (columns) are labelled by their cluster and cell cycle phase. MG(3A) upregulates neuroprotective genes while MG(3B) upregulates many of the neurodegerative genes. Gene lists are separated by the solid black line. Note that *Smad3* was not in the HVGs. C) Boxplot of the median expression of neuroprotective genes across the microglia clusters. Cells are binned by Fabp5 expression levels. High > 15 counts; 1 < Med < 15; Low < 1 (note the term "none" was not used to account for sparsity). *Fabp5*$^{+}$ cells express lower levels of neuroprotective genes.

To summarise, this myeloid cell atlas suggests that acute phase microglia upregulate *Msr1* and NF-κB signalling, become pro-inflammatory, and undergo proliferation, adjacent to the lesion core and in the NASC (Figure 5.25). Some of these proliferating cells transition into a DAM1-like state of phagocytosis, but it is unclear if all the proliferating cells adopt this state. In the late acute and subacute phase, microglia drastically upregulate the lipid-processing genes *Trem2* and *Apoe* and become cytotoxic, producing ROS and IL-6. A subset of microglia remain in this state indefinitely. Others transition into a neuroprotective, homeostatic-like state. A final subset takes on a lipid-processing, foam cell-like, scar formation-promoting state, that shares a transcriptional signature with the previously described MGnD and is found in both the NASC and lesioned areas. The lipid processing state is also adopted by a population of subacute and chronic monocyte-derived macrophages. Beyond the early chronic (21 dpi) timepoint, it is unclear whether these states persist, or if cells transition into other, unidentified states.

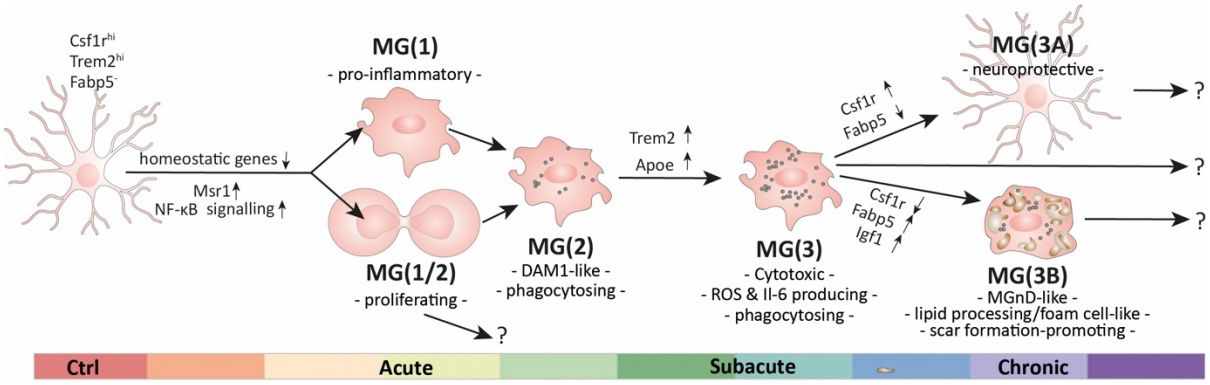


Figure 5.25 A summary of the microglia states present in the time-resolved myeloid cell atlas across the acute, subacute, and early chronic phases of SCI.

5.5.3 Characterising Infiltrating Myeloid Clusters

In contrast to the microglia clusters, which were well-represented across the time-resolved atlas, infiltrating myeloid clusters were dominated by cells from 1-3 dpi (Figure 5.26). At 1 dpi, three infiltrating myeloid cell types were present: neutrophils, T-Cells, and monocyte-derived cells, in line with previous findings (Alizadeh et al., 2019; Donnelly & Popovich, 2008; Milich et al., 2021).

Neutrophils were present only in the acute phase of SCI, showed little evidence of cell cycling (Figure 5.26B), and presented the expected GO terms of phagocytosis, neutrophil aggregation, and intrinsic apoptosis (Zivkovic et al., 2021). T-Cells were present across the SCI pathology (Figure 5.26A) as expected (Donnelly & Popovich, 2008), however given that only 164 T-Cells were collected, I was not confident that I could draw any meaningful conclusions from their transcriptional profiles.

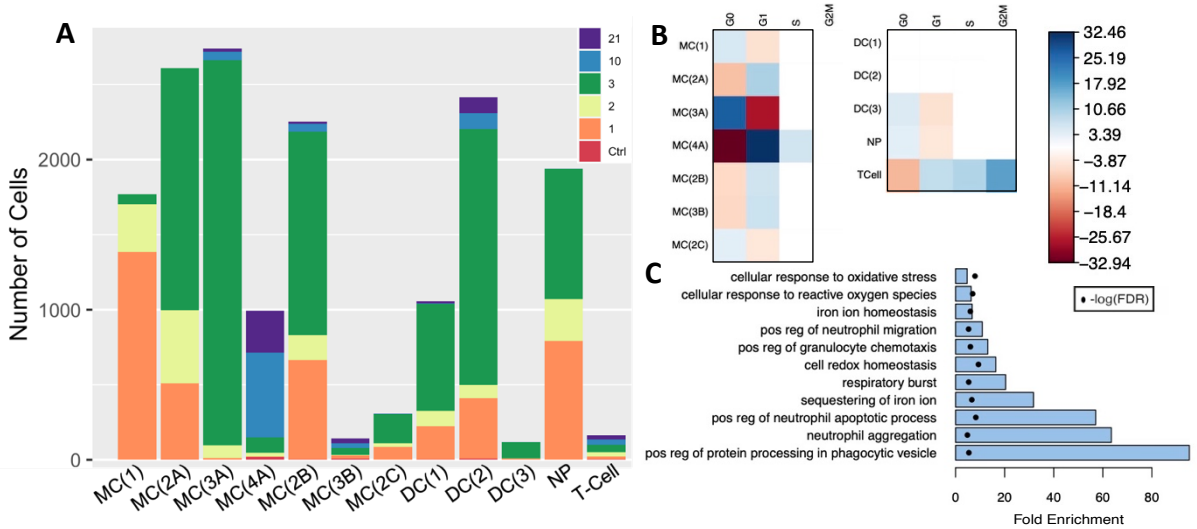
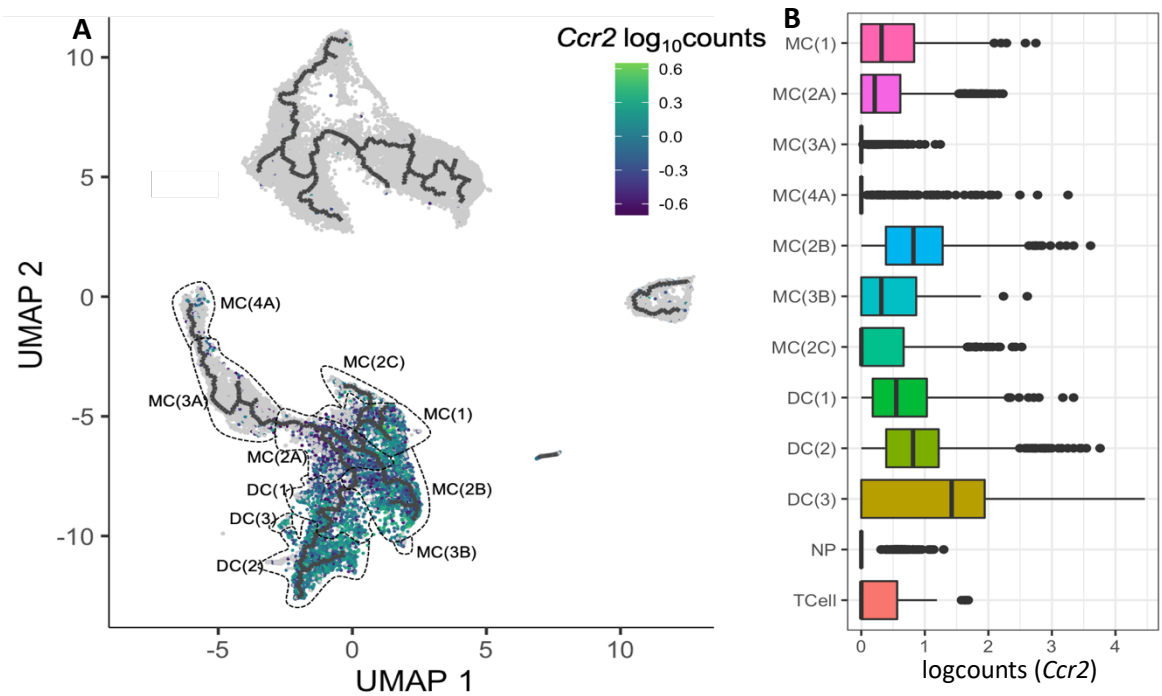


Figure 5.26 A) Bar chart of the infiltrating clusters depicting the number of cells per cluster, coloured by their dpi. B) A correlation plot of the Pearson's chi-squared residuals between infiltrating cluster membership (rows) and cycle (columns). Pearson's Chi-squared test X-squared = 2414, df = 33, p-value < 2.2e-16. C) The fold enrichment and FDR of select GO terms for the NP cluster. Pos = positive; Neg=negative, reg = regulation. The expected GO terms for neutrophils are present.

Monocyte-derived cells were more plentiful and dynamic than neutrophils and T-Cells. Cluster MC(1) comprised infiltrating cells from mainly 1 dpi. Based on Veronica's quantification (Figure 4.5) and the expression of C-C Motif Chemokine Receptor 2 (*Ccr2*) (Phillips et al., 2005), MC(1) had yet to fully infiltrate the spinal cord parenchyma and differentiate into macrophages or dendritic cells, as previously reported for 1 dpi (Alizadeh et al., 2019; Donnelly & Popovich, 2008; Milich et al., 2021) (Figure 5.27A, B). This cluster was characterised by GO terms for the response to low oxygen levels and angiogenesis and the initiation of fibrosis, both of which were previously reported to stem from macrophages after 3 dpi (Kigerl et al., 2009; Whetstone et al., 2003). This discrepancy might be a consequence of transcription *vs* protein-level analyses. MC(1) was also enriched with GO terms for the production and response to cytokines and reactive oxygen species (ROS), oxidative stress-induced neuronal death, respiratory bursts, phagocytosis, cellular detoxification, blebbing, and lipid processing (Figure 5.27C). *Fabp5* was expressed in MC(1) cells, supporting a lipid processing function (Zhu et al., 2017) (Figure 5.27D).



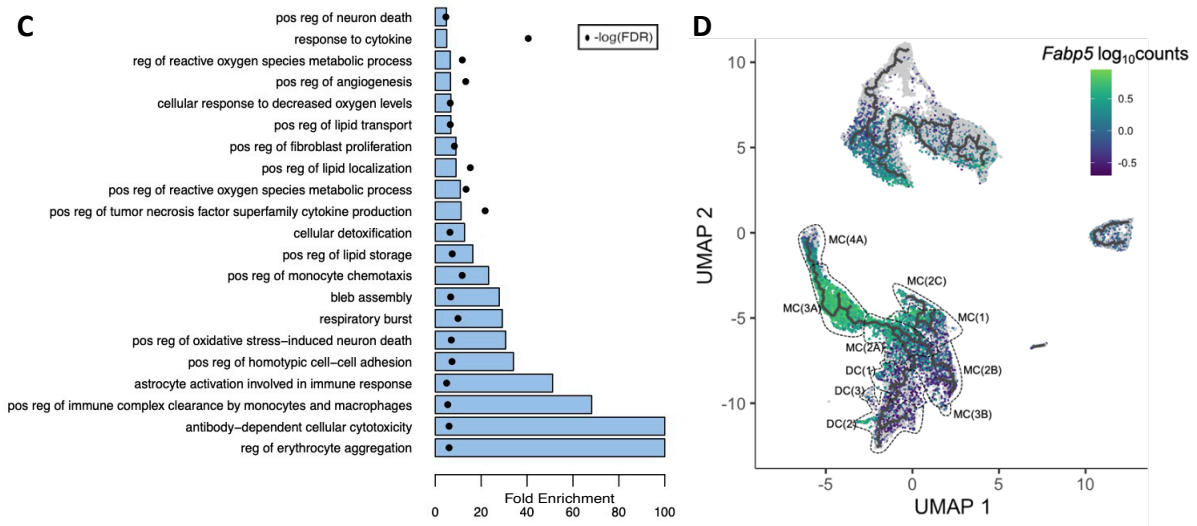


Figure 5.27 A) UMAP coloured by the log_{10}counts of *Ccr2*. The UMAP is superimposed with the trajectory inferred via *Monocle 3*. Infiltrating clusters are denoted by the dotted lines. *Ccr2* is differentially expressed across the infiltrating myeloid cell atlas, with the highest expression in the acute phase monocytes and the monocyte-derived dendritic cells. B) Boxplots quantifying the *Ccr2* the expression shown in A). Note these are the log_2 normalised counts. C) The fold enrichment and FDR of select GO terms for the MC(1) cluster. Pos = positive; Neg=negative; reg = regulation. D) UMAP coloured by the log_{10}counts of *Fabp5.*

From MC(1), monocyte-derived cells projected down three main trajectories, which I arbitrarily labelled from A-C. Starting with the shortest trajectory, MC(2C) comprised cells from 1-3 dpi (Figure 5.26A), upregulated *Fabp5* (Figure 5.27D), and expressed GO terms for IL-4 signalling, which has been shown to be neuroprotective in the context of SCI (Fenn et al., 2014). Terms relating to phagocytosis, and blebbing were also enriched (Figure 5.28A). Furthermore, this cluster was uniquely defined by an increase in the expression of *Lrp1* (Figure 5.28B), an endocytic receptor involve in phagocytosis of apoptotic cells (May et al., 2002). Thus, MC(2C) may represent a small subset of infiltrating myeloid cells that are activated by IL-4 signalling and clear apoptotic cells after SCI before undergoing apoptosis themselves.

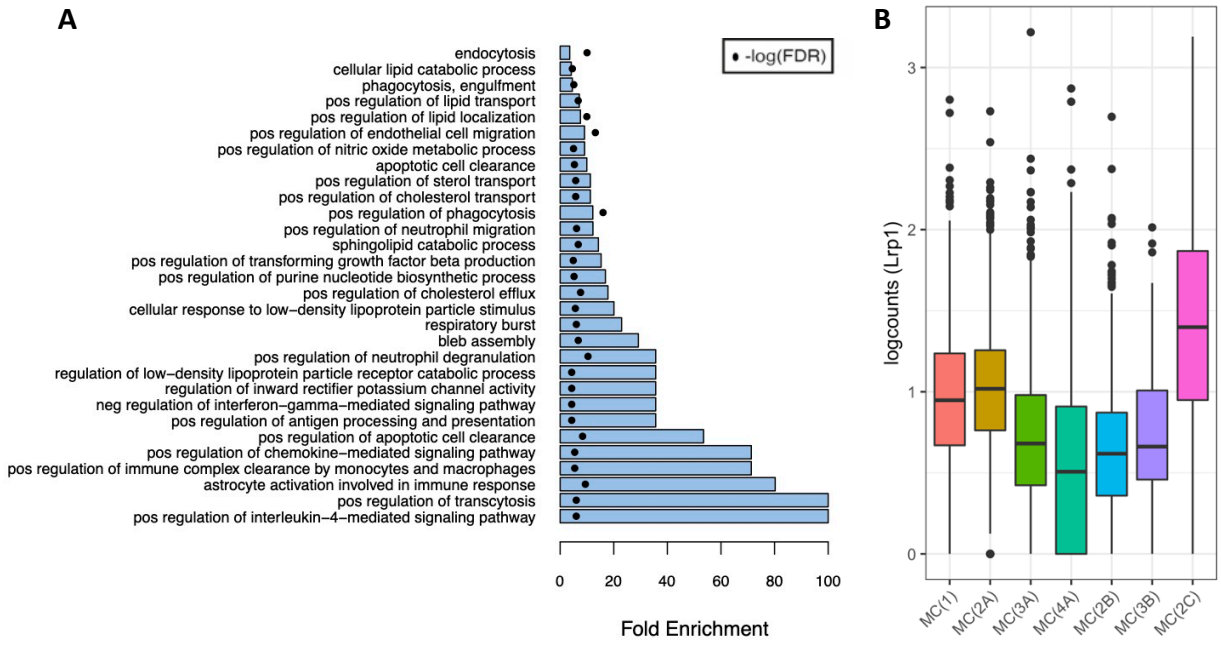


Figure 5.28 The fold enrichment and FDR of select GO terms for the MC(2C) cluster. Pos = positive; Neg=negative. B) Boxplots quantifying the expression of *Lrp1,* which is starkly upregulated in cluster MC(2C).

MC(2B) and (3B) were *Ccr2*-expressing (Figure 5.27A,B), acute phase clusters with minimal contributions from 10 and 21 dpi (< 2.5%) (Figure 5.26A) and no strong evidence of cell cycling (Figure 5.26B). These cells were characterised by GO terms for blebbing and seemed to be fated towards cytokine production and apoptosis as described (Greenhalgh & David, 2014) (Figure 5.29A). However, these clusters also contained GO terms for the negative regulation of MyD88 signalling, a pro-inflammatory, NF-κB-activating, macrophage-recruiting signalling pathway that has been demonstrated to have degenerative effects in rat models of compression SCI (Boivin et al., 2007; Xu et al., 2018; Yao et al., 2012). MC(2B) also showed a positive regulation of neuroprotective IL-4 signalling. Notably, this potentially neuroprotective cluster of infiltrating macrophages expressed very low levels of *Fabp5* compared to other infiltrating clusters (Figure 5.29B; Figure 5.27D), supporting the hypothesis that *Fabp5* expression may have negative effects in the SCI pathology.

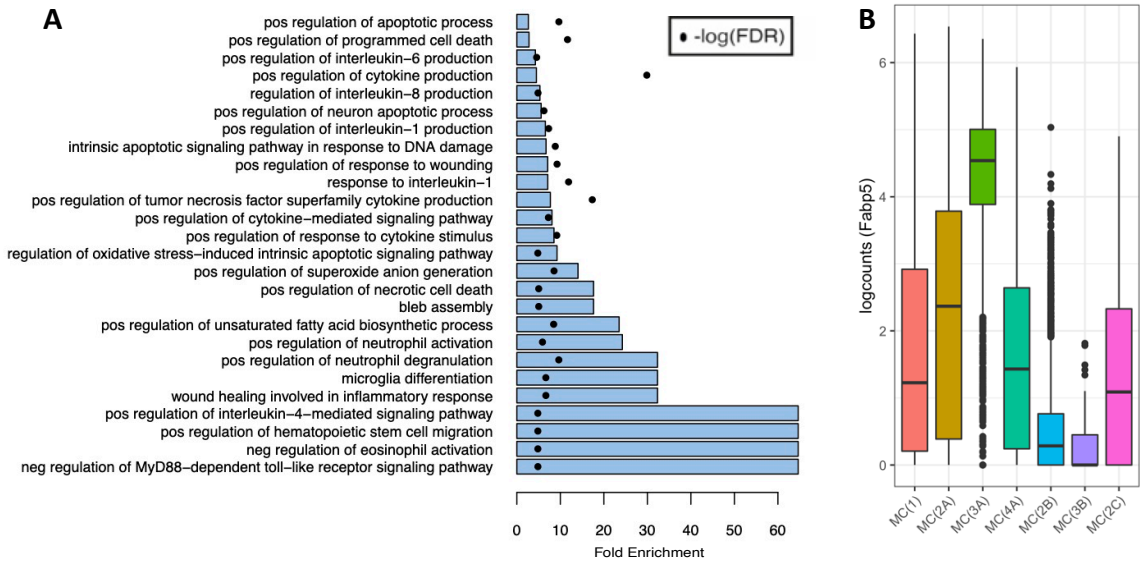


Figure 5.29 The fold enrichment and FDR of select GO terms for the MC(2B) cluster. Pos = positive; Neg=negative. B) Boxplots quantifying the expression of *Fabp5,* which is lowly expressed in clusters MC(2B) and MC(3B) compared to other infiltrating monocyte-derived clusters.

MC(2A) was contributed to by infiltrating cells from 1-2 dpi but dominated by 3 dpi (Figure 5.26A). The cluster was characterised by GO terms for the negative regulation of intrinsic apoptosis, response to hypoxia, and NF-κB transcription factor activity (Figure 5.30A). Notably, NF-κB is upstream of *Fabp5*, which was drastically upregulated in MC(3A) (Figure 5.29B) (Kaczocha et al., 2014), the next stage in the *Monocle3* predicted trajectory (Figure 5.27D). In fact, the fate of MC(2A) correlated with its *Fabp5* expression, with *Fabp5*^{lo} cells differentiating into pro-regenerative dendritic cells (Yaguchi et al., 2009), whereas *Fabp5*^{hi} cells became pro-inflammatory macrophages (Figure 5.27D). By binning cells based on *Cd74* expression, which is highly upregulated in dendritic cells, or *Ccr2* which is downregulated upon macrophage differentiation (Phillips et al., 2005), I saw that both high *Cd74* and *Ccr2* expression correlated with low *Fabp5* expression, even as early as MC(1) (Figure 5.30B). In further support of the trajectory analysis, the trajectory-predicted fate of MC(2A) into dendritic cells or macrophages was well supported by the GO analysis, which returned terms for myeloid cell differentiation and purine metabolism (Ahmed & Weidemann, 1994) (Figure 5.30A).

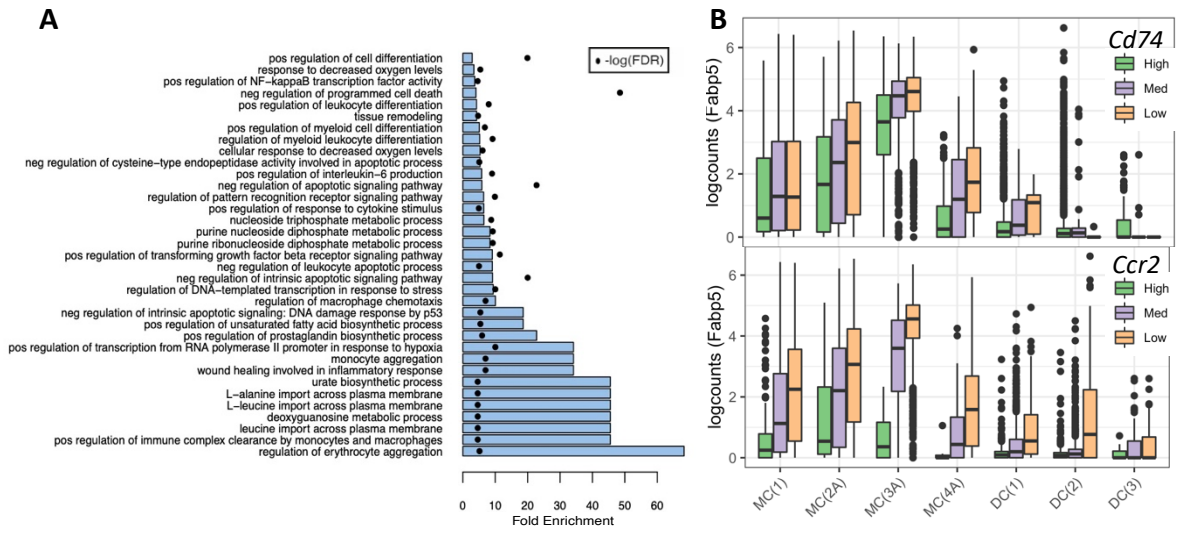
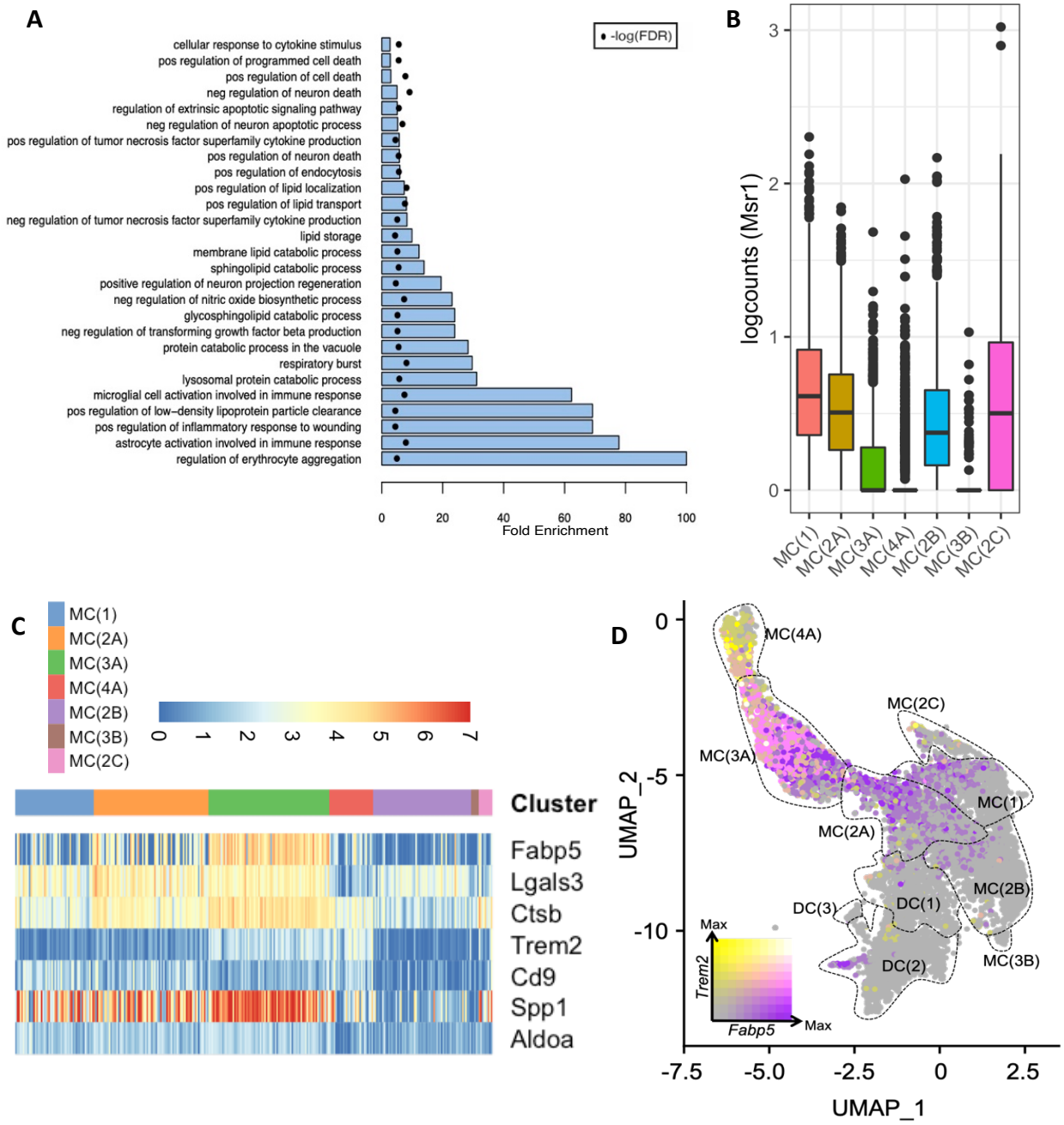


Figure 5.30 A) The fold enrichment and FDR of select GO terms for the MC(2A) cluster. Pos = positive; Neg=negative. B) Boxplot of the *Fabp5* expression across the MC(1), MC(A), and DC clusters. Top: Cells are binned by *Cd74* expression levels. High >= 20 counts; 1 < Med < 20; Low =< 1. Bottom: Cells are binned by *Ccr2* expression levels. High >= 10 counts; 1 < Med < 10; Low =< 1 (note the term "none" was not used to account for sparsity). *Cd74*$^{+}$ cells express lower levels of *Fabp5* while *Ms4a7*$^{+}$ cells express higher levels of *Fabp5*.

The monocyte-derived dendritic cell clusters were dominated by the acute phase, but a portion of cells remained present at 10 and 21 dpi (Figure 5.26A). This possibly occurred through local proliferation (Swirski et al., 2014) as a small portion of DC(2) and (3) cells were assigned to the G1 phase (Figure 5.7B) and the GO analysis returned terms for G1/S transition. The dendritic cell clusters were also dominated by the expression of MHC II genes and GO terms for antigen presentation, phagocytosis, and T-cell chemotaxis, as expected (Guilliams et al., 2014).

The *Ccr2*$^{-}$ monocyte-derived macrophage clusters, MC(3A) and MC(4A) (Figure 5.27A,B), comprised 3 dpi and 10/21 dpi cells, respectively (Figure 5.26A), in line with previously described timelines for macrophage infiltration post-SCI (Alizadeh et al., 2019; Donnelly & Popovich, 2008; Greenhalgh & David, 2014; Milich et al., 2021). MC(3A) was highly correlated with the G0 phase (Figure 5.26B) and was characterised by a uniform upregulation of *Fabp5* (Figure 5.29B). In line with this observation, GO term analysis showed highly enriched terms for lipid metabolism and processing phagocytic cargo, but terms for phagocytosis were absent (Figure 5.31A). These features closely resembled cluster MG(3B), thus I investigated whether MC(3A) might also acquire a foam-cell profile and found that it did, including the upregulation of *Msr1* in the preceding clusters (Figure 5.31B,C) (Kong et al., 2020).

Other enriched GO terms in MC(3A) included microglia and astrocyte activation, neurite growth and wound healing but also the mediation of extrinsic apoptosis, and both the positive and negative regulation of nitric oxide and cytokine production (Figure 5.31A). These opposing terms could be explained by the transition within MC(3A) from a predominantly lipid processing, *Fabp5*^{hi}, 3 dpi state to a *Trem2*^{hi} MG(4A)-like state at 10-21 dpi (Figure 5.31D). We verified this dynamic expression pattern via smFISH (Figure 5.31E), as described (4.2.2).



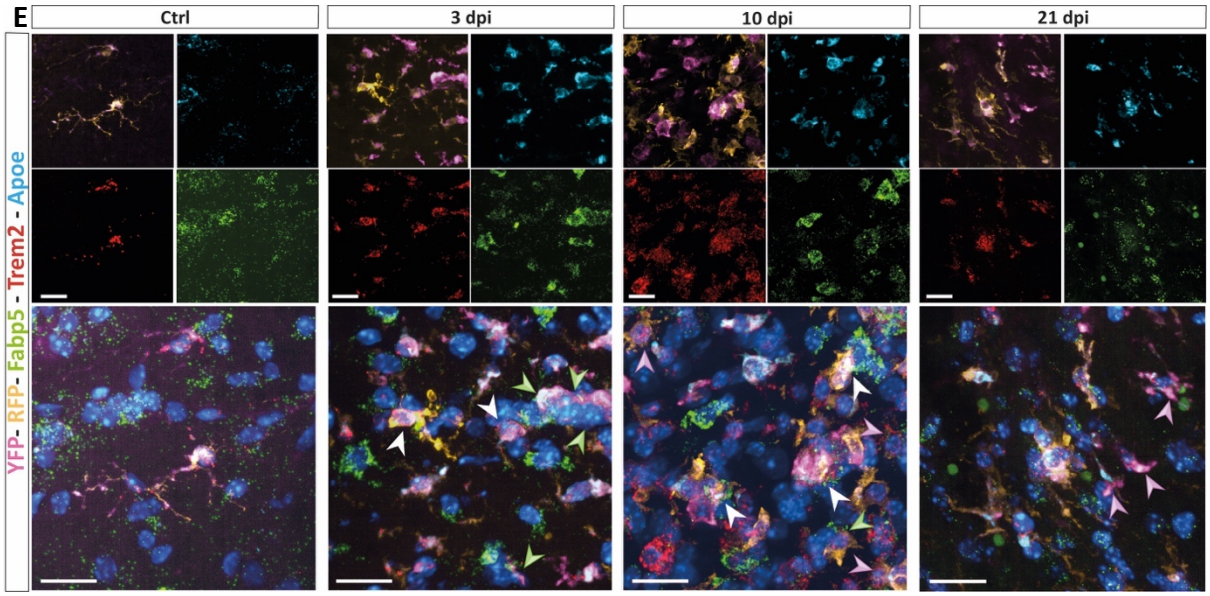


Figure 5.31 A) The fold enrichment and FDR of select GO terms for the MC(3A) cluster. Pos = positive; Neg=negative. B) Boxplot of *Msr1* expression in the MC clusters. *Msr1* is upregulated before *Fabp5* and the adoption of the foam cell like phenotypes. C) Heatmap of *Fabp5* and foam cell-associated genes in the MC clusters. MC(3A) upregulates *Fabp5* and foam cell-associated genes, similar to MG(3B). D) UMAP of infiltrating monocyte-derived cells coloured by the scaled logcounts of *Fabp5* and *Trem2*. Note the transition within MC(3A) from *Fabp5*^{hi} to *Trem2*^{hi}. Clusters are denoted by the dotted lines. E) smFISH depicting the expression of *Fabp5*, *Trem2* and *Apoe* in the lesion epicentre at Ctrl, 3, 10, and 21 dpi. Nuclei in the merged panels were stained with DAPI. Arrowheads are as follows: White = Fabp5^{+}/RFP^{+}/YFP^{+} resident myeloid cells; Green = Fabp5^{+}/Trem2^{-}/RFP^{-}/YFP^{+} infiltrating myeloid cells; Purple = Fabp5^{-}/Trem2^{+}/RFP^{-}/YFP^{+} infiltrating myeloid cells. Scale bars: 20 µm.

Unlike MC(3A), the 10-21 dpi MC(4A) was highly correlated with the G1 cell cycle phase (Figure 5.26A,B). This suggests these infiltrated macrophages underwent local proliferation in the subacute and chronic phases of SCI, which has been proposed (Greenhalgh & David, 2014; Swirski et al., 2014). Surprisingly, GO analysis returned many of the same terms as MG(3) (Figure 5.19A): glia activation, antigen presentation, cytokine production and monocyte chemotaxis, ROS-producing reverse electron transport (Peruzzotti-Jametti et al., 2021), cell killing, and inflammation-perpetuating IL-6 production (Hunter & Jones, 2015). Furthermore, MC(4A) comprised not only infiltrating macrophages, but also a population of CAMs, distinguished by a unique expression pattern and the resident myeloid cell fate mapping label (Figure 5.32B,C). This population was likely driving the antigen presentation GO terms.

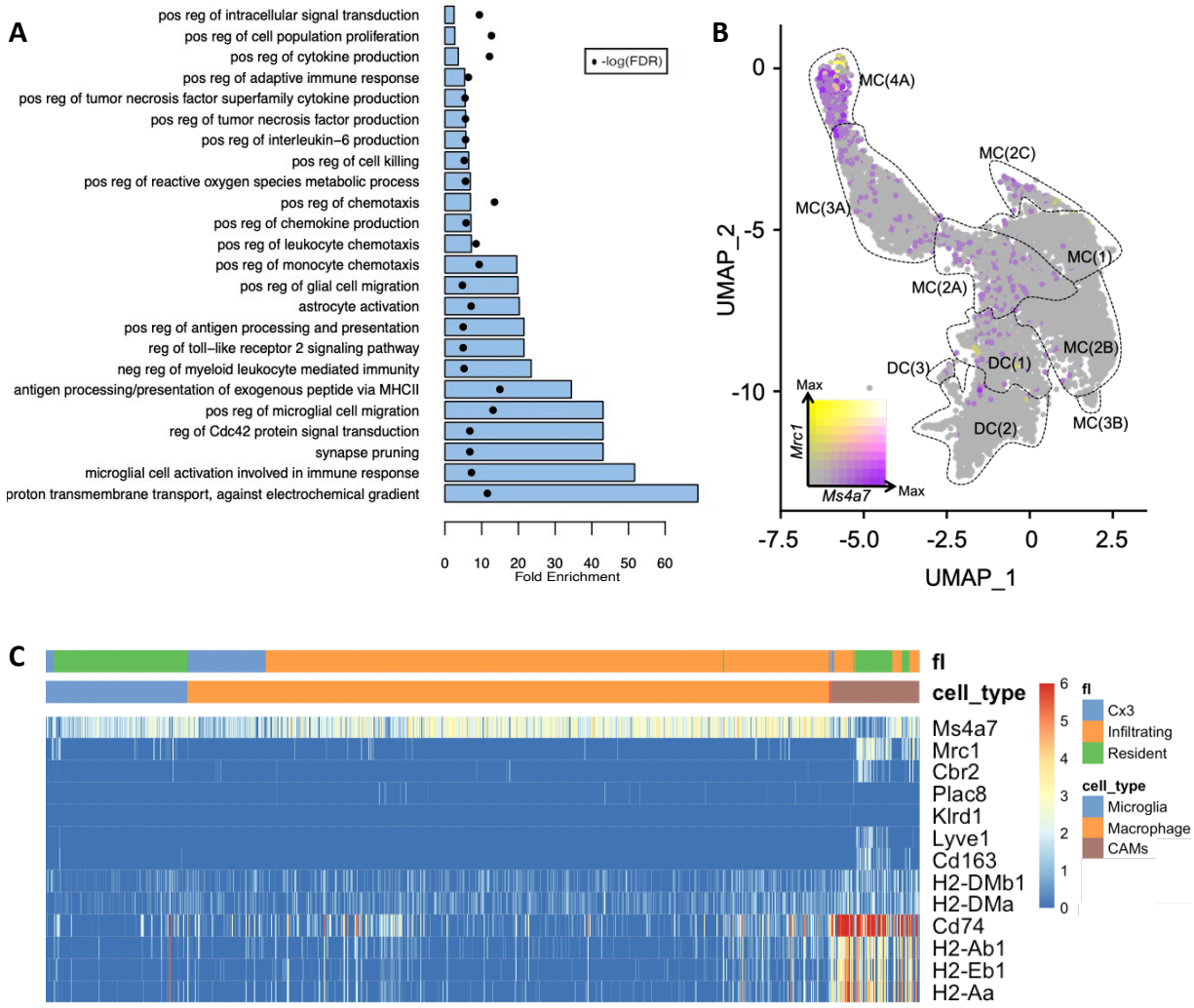


Figure 5.32 A) The fold enrichment and FDR of select GO terms for the MC(4A) cluster. Pos = positive; Neg=negative; reg = regulation. B) UMAP of infiltrating monocyte-derived cells coloured by the scaled logcounts of *Mrc1* and *Ms4a7*. CAMs upregulate *Mrc1* compared to other MC(4A) cells. Clusters are denoted by the dotted lines. C) A heatmap of CAM marker expression in cluster MC(4A). Cells (columns) are labelled by their fate mapping status (fl) and annotated cell type. The CAMs have a distinct transcriptional profile and were found predominantly in the RFP⁺/YFP⁺ resident samples.

The smFISH experiment described in *Characterising Microglia Clusters* was designed specifically for the microglia clusters. However, the monocyte-derived A trajectory shared the dynamic expression patterns of *Apoe*, *Trem2,* and *Fabp5*. Thus, I loosely defined each cluster according to Figure 5.33A,B, plus the fate mapping labels. Given that the cluster definitions were not perfectly aligned, I used this data to draw general conclusions (Figure 5.33C). At 3 and 10 dpi, infiltrated myeloid cells were present across the spinal cord. By 21 dpi, very few infiltrated cells were observed, in contrast to Veronica's quantifications (Figure 4.6) and previous reports, but the detected infiltrated cells were in the lesion core as expected (Alizadeh et al., 2019).

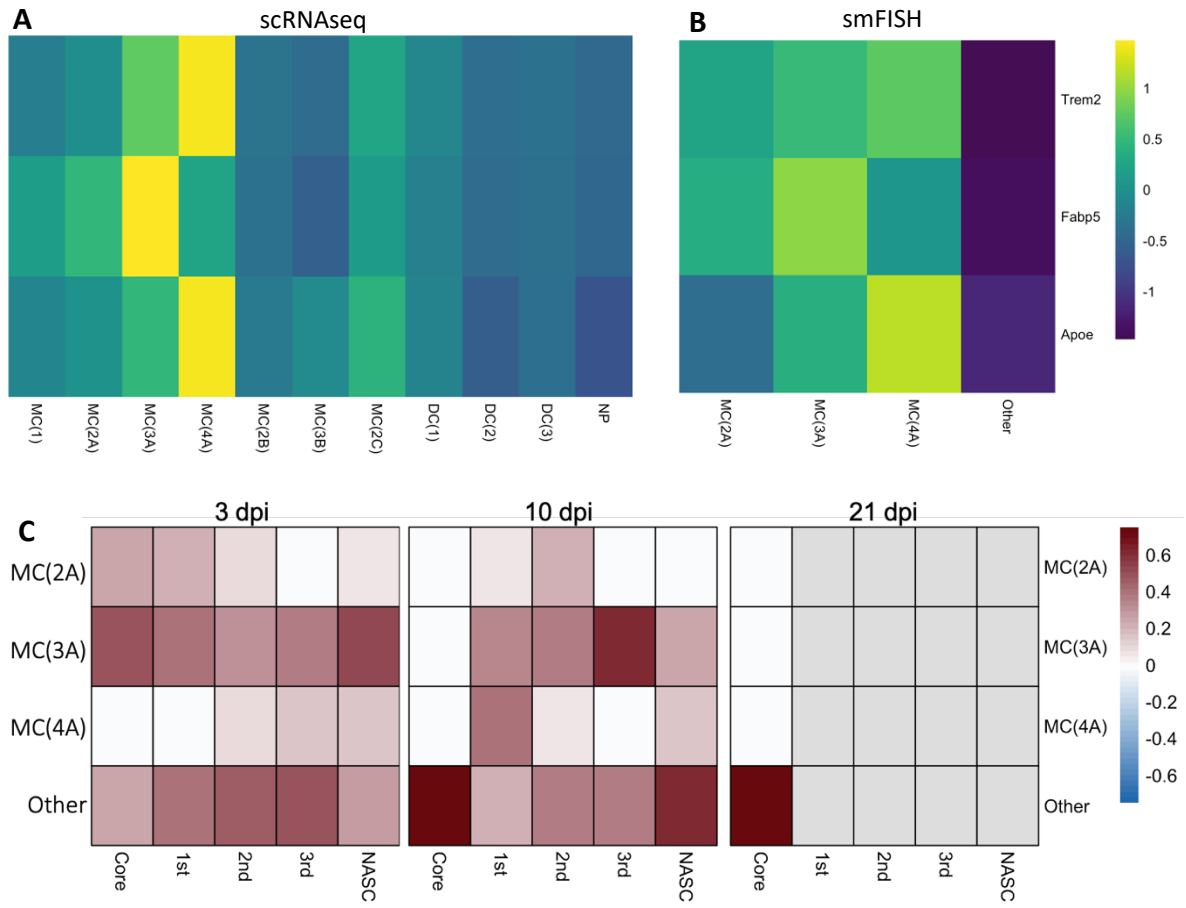


Figure 5.33 A) A heatmap of the scRNAseq data showing the average normalised expression of each gene of interest for each infiltrating myeloid cell cluster, scaled across each gene (row). B) A heatmap of the smFISH data depicting the average number of spots detected per cluster. The cluster definitions used for smFISH are not particularly well matched as they were designed for the microglia clusters. D) The smFISH-based quantification of MC(A) clusters by ROI at several dpi, presented as normalised residuals (condition – control/the number of microglia in the ROI).

Finally, I compared the monocyte-derived clusters with those from Milich et al., 2021. I found that the expression patterns of DC(1) closely matched their DC profiles (Figure 5.34), as did MC(1) and their monocytes. Their inflammatory macrophage cluster was well-aligned with the MG(4A) cluster, in dpi (predominantly subacute), marker gene expression (Figure 5.34) and GO terms for glia activation (Figure 5.19A). This suggests that MG(4A) appears as early as 7 dpi. Their chemotaxis-inducing macrophage cluster was similar to clusters MC(1), (2A), and (2C), in gene expression, GO terms for chemotaxis, and dpi (predominantly acute). Notably, MG(3A) and (4A) were also enriched for GO terms relating to chemotaxis, despite not matching the proposed "chemotaxis-inducing" profile. Finally, while I observed proliferation in the MC(4A) cluster, it did not match their dividing myeloid cell cluster and I did not observe an upregulation of *Cdk1* in any of the infiltrating myeloid cell clusters (see *5.6 Discussion*).

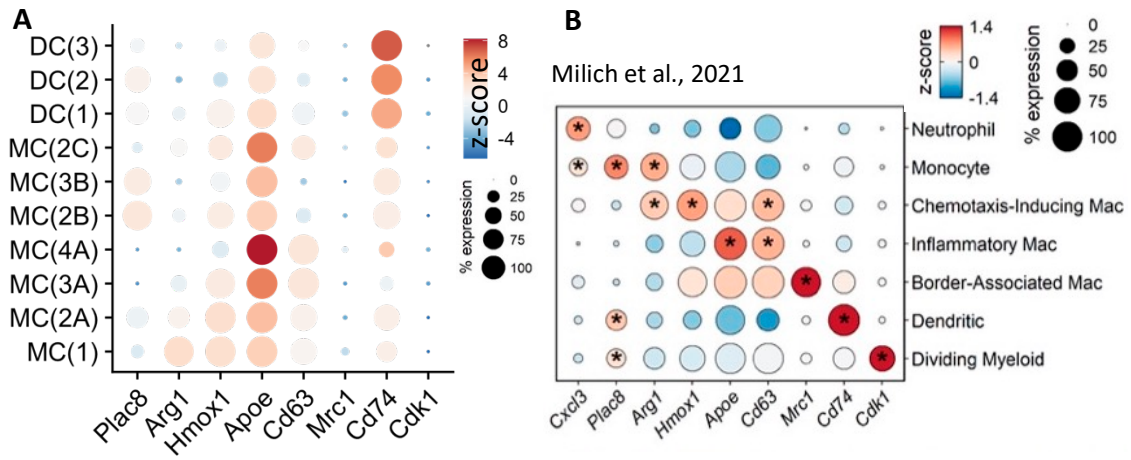


Figure 5.34 Dot plot of the expression of genes identified by Milich et al., 2021 as distinguishing infiltrating myeloid cell types across the clusters. The expression is shown for monocyte-derived clusters in this myeloid cell atlas (A) and Figure 4C from Milich et al., 2021 (B).

In summary, after SCI, monocytes infiltrate the lesioned spinal cord where they upregulate *Msr1* and adopt a phagocytic, ROS and cytokine-producing phenotype as early as 1 dpi. These monocytes transition into one of three states: an IL-4 signalling, apoptotic debris-clearing state that ends in intrinsic apoptosis; a cytokine-producing but MyD88 signal attenuating state that may play a neuroprotective role before apoptosis; or a state of differentiation and NF-κB transcription factor activity. The latter persists into the subacute and chronic phases as either dendritic cells or macrophages, and this bifurcation is correlated with *Fabp5* and *Ccr2* expression. The macrophages adopted a lipid metabolising, glia-activating, foam cell-like phenotype in the acute phase, but by the subacute and chronic phases had transitioned to a cytotoxic, ROS and IL-6 producing state with parallels to the MG(3) cluster. Furthermore, this cluster included a small population of CAMs, suggesting they have a similar transcriptional profile and potentially a similar role at this stage in the pathology. Both dendritic cells and macrophages persisted to 21 dpi and, as with the microglia, their fate beyond this point is unclear.

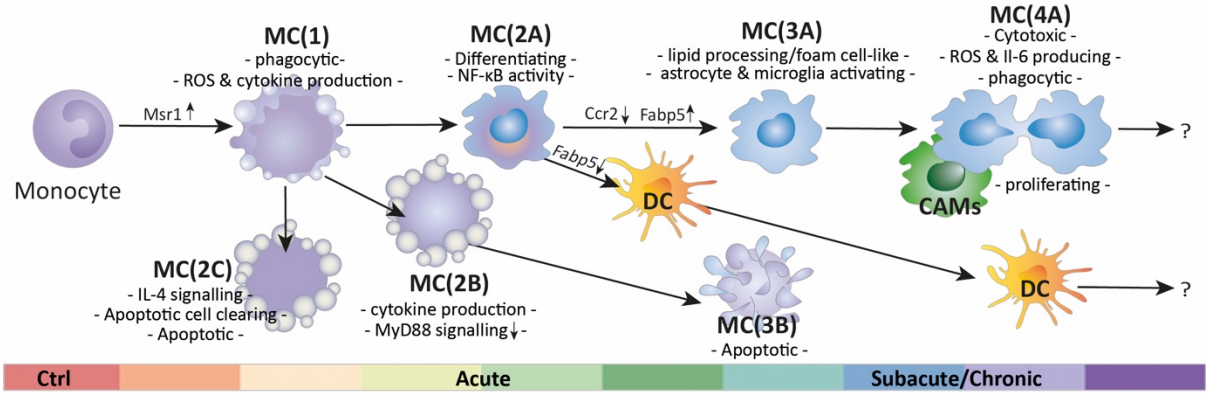


Figure 5.35 A summary of the monocyte-derived states, and CAMs, present in the time-resolved myeloid cell atlas across the acute, subacute, and early chronic phases of SCI.

5.6 DISCUSSION

In this chapter I have demonstrated that this time-resolved SCI myeloid cell atlas exhibits the expected qualities based on previous findings. I then analysed this dataset in its own right by first performing unsupervised clustering and trajectory analysis and then characterising each cluster to finally generate a time-resolved map of transcriptional profiles across the SCI pathology.

The goal of a typical study is to uncover a new phenomenon. However, demonstrating that previous findings are replicable, especially through different protocols or even in new disease contexts, is equally critical (Nosek & Errington, 2020). Using this novel myeloid cell atlas, I investigated the expression of canonical microglia markers upon activation, and found that they were downregulated, as increasingly reported under many other pathological conditions including, very recently, SCI (Bennett et al., 2018; Butovsky et al., 2014; Dubbelaar et al., 2018; Hammond et al., 2019; Jordão et al., 2019; Milich et al., 2021). This downregulation is now accepted as a first step in a disease-independent microglial activation program, termed DAM1 (Keren-Shaul et al., 2017). After SCI, I found that microglia began to adopt a DAM1-like phenotype in the MG(1) state as early as 1 dpi. *Apoe* and *Lyz2,* however*,* were only robustly upregulated from 2 dpi onwards (Figure 5.5). Additionally, unlike the DAM1 microglia described in the context of AD and ALS (Keren-Shaul et al., 2017), these DAM1-like cells were proliferating. Despite the proliferation, and the drastically different timescales between this MG(1) cluster and the DAM1 microglia associated with the progression of AD and ALS, the DAM1 profile was well replicated in this SCI myeloid cell atlas. Unlike the DAM1 profile, DAM2 genes were only partially upregulated in microglia by the subacute and chronic phase of SCI, with *Axl*, *Csf1*, *Lilrb4a,* and *Itgax* noticeably absent (Figure 5.5). Perhaps at a more chronic time point the DAM2 profile would be recapitulated in SCI, or this discrepancy may represent a level of pathology-specificity in the DAM phenotype.

DAM2 is hypothesised to be neuroprotective in the contexts of AD and ALS (Keren-Shaul et al., 2017), but in this myeloid cell atlas, it was inversely related to the neuroprotective M0

profile (Figure 5.15B; Figure 5.24B) (Krasemann et al., 2017) and, in the case of cluster MG(3B), overlapped with the neurodegenerative MGnD profile (Krasemann et al., 2017). The other DAM2-like cluster, MG(3), was characterised by the upregulation of most MGnD genes, and neurodegenerative GO terms including ROS and IL-6 production, and cytotoxicity. Thus, there is little evidence that the DAM2-like profile is neuroprotective in the context of SCI. Notably, the MGnD phenotype, like DAM2, did not have a perfect match in the SCI atlas, with *Itgax* and *Ccr2* expression lacking. The MGnD and DAM profiles were characterised via bulk RNAseq and MARS-seq (Jaitin et al., 2014), respectively, both of which provide greater sensitivity compared to the 10X Chromium platform. Thus, if these genes were lowly expressed in the cells, their absence in the dataset could be due to the relatively low sensitivity of the platform.

Very recently, CNS resident and infiltrating cells from 8 mm of SCI tissue were analysed at 1, 3, and 7 dpi in a contusion model of SCI (Milich et al., 2021). This study used female adult WT mice (65-kdyne impact to T8) and the 10X Chromium scRNAseq platform with version 2-3 chemistry. Milich et al., made no reference to the DAM or MGnD phenotypes. They identified four microglia clusters, which they annotated based on GO term analysis: homeostatic, dividing, inflammatory, and migrating. The "dividing" cluster was most prominent at 1 dpi, in line with my findings. A portion of cycling cells were present at 7 dpi, which might be analogous to the 10-21 dpi MG(2) cells. Their "inflammatory microglia" cluster was defined as *P2ry12*^{lo}*/Igf1*$^{+}$. In this myeloid cell atlas, that would correspond to MG(3) and (3B) (Figure 5.19B). Many of these cells were also *Ms4a7*$^{+}$ and *Siglech*$^{-}$ and based on the fate-mapping results in this atlas, were likely infiltrating macrophages mis-annotated as microglia. Like MG(3), this "inflammatory microglia" cluster was characterised by GO terms for extrinsic cytotoxicity and cytokine production. However, Milich et al., observed this phenotype as early as 1 dpi. The time-resolved expression of *Igf1* was not provided, thus it's unclear whether the datasets are truly conflicting. Their "migrating" cluster was present in small proportions across all three dpi and were characterised by GO terms for migration and motility and as *P2ry12*^{lo}*/Igf1*$^{+}$*/Msr1*$^{+}$. I observed these GO terms in each of the microglia clusters and did not detect the co-expression of *Igf1* and *Msr1* in any myeloid cells. Milich et al., processed their data using Seurat v3 and, their code is not available, but based on their methods section, it appears they used MNN to merge each sample, irrespective of dpi. Thus, it's possible that this

cluster is the combination of several cell states. The 10-21 dpi MG(3A) state is absent from their dataset, in line with their collection timepoints.

Milich et al., also captured infiltrating myeloid cells in their dataset including CAMs, neutrophils, dendritic cells, monocytes, and macrophages. At 1 dpi, they found a population of *Cdk1*$^{+}$ cycling monocyte-derived cells, which I did not observe. Notably, this cluster comprised cells from each dpi and the Ctrl, suggesting this was an isolation-induced artefact. It likely stemmed from their enzymatic dissociation approach during which the dissected tissue was left at 37° for 30 minutes, a procedure that has been shown to activate microglia, and likely affects other myeloid cells (Hammond et al., 2019). Milich et al. also observed neutrophils, monocytes, and a subset of macrophages that they termed, "chemotaxis-inducing." These monocytes and macrophages still expressed *Ccr2*, in line with my findings. This cluster was similar to MC(1), (2A), and (2C), in gene expression and GO terms for chemotaxis, suggesting that a level of heterogeneity was overlooked in their dataset. This might have been due to the suspected over-correction of their samples.

By 7 dpi, Milich et al., found dendritic cells, *Ms4a7*$^{+}$ "chemotaxis-inducing" macrophages and "inflammatory macrophages", in line with this myeloid cell atlas. Their "inflammatory macrophage" cluster largely aligned with MC(4A), suggesting it appears as early as 7 dpi. They also observed *Ms4a7*$^{+}$*/Cd74*$^{+}$*/Mrc1*$^{+}$ CAMs, which clustered closely to their "inflammatory macrophage," just as I observed with MC(4A) in this dataset. The Milich et al., "chemotaxis-inducing" macrophage cluster was similar to 3-21 dpi cells from MC(3A), with the exception of *Arg1*, which was not observed beyond 3 dpi. This might simply reflect the difference between 7 and 10 dpi, however chemotaxis-related GO terms were observed in 10 and 21 dpi MC(3A) and (4A) cells, indicating that their proposed "chemotaxis-inducing" signature is not comprehensive.

In summary, the Milich et al., 2021 myeloid cell data is generally well-aligned with this myeloid cell atlas, particularly regarding the dynamic presence or absence of cell types across the SCI pathology. However, there are some noticeable differences that appear to be caused by the contrasting isolation protocols and the computational analyses. The latter exemplifies the

importance of making analysis scripts publicly available, to ensure reproducibility and allow for more detailed comparison across datasets.

To investigate and characterise this time-resolved myeloid cell SCI atlas, beyond the comparison with previously described gene lists, I relied heavily on unsupervised clustering and trajectory analysis. Unsupervised clustering is a powerful tool that allows us to assign biological meaning to groups of transcriptionally-similar cells. The clustering approach used in this chapter, while efficient and boasted to return well-connected communities (Traag et al., 2019), suffers from density-bias, where more numerous populations result in more clusters, due to the density of the graph and not necessarily reflecting an increase in heterogeneity (Amezquita et al., 2020). This clustering approach also requires a user-defined number of nearest neighbours (k) for the k nearest neighbour graph. The value of k alters the resolution of the clusters, with a higher k value producing larger clusters that contain more substructure, and vice versa. Ultimately, this resolution flexibility has important implications for DEG and GO analysis.

This study relied heavily on GO enrichment analysis of DEGs to infer putative functions for clusters of cells. GO is a community-curated resource that provides functional annotations to gene products, which are then grouped together based on shared biological properties ("Gene Ontology Consortium: going forward," 2015; Young et al., 2010). This resource can then be used to identify GO categories that are overrepresented in a DEG list. This approach has several caveats. First, the result of GO analysis is constructed from the DEG list provided. Thus, any error or bias in DEG analysis, and consequently clustering and pre-processing, will be reflected in the GO analysis. Comparing the cluster-based GO terms with relevant studies, if available, or replicating one's own analysis with diverse workflows can help to identify more robust and possibly more relevant terms. Second, a GO term can oversimplify or even misdirect the biological interpretation of the cluster because the effect of a gene can be cell type- or pathology-dependent or have multiple functions (E Hirbec et al., 2017; Reimand et al., 2019). This highlights the need to cross-check enriched GO terms with other cluster properties, for e.g., cell cycle analysis or non-GO curated gene lists, whenever feasible. Finally, GO analysis infers protein-level function based on mRNA transcripts, which is a general limitation of scRNAseq that will be discussed in *6.1*. GO terms and scRNAseq analysis in

general can act as a powerful descriptive guide and can insinuate function on a single cell level, but ultimately follow-up functional studies are required to clearly demonstrate any causal effects.

Another tool that was heavily exploited in this chapter was trajectory analysis. This is a useful approach for identifying transitions between cell states. Paired with the time-resolved labels in this dataset, trajectory analysis allowed me to interpret the transition of myeloid cell states across the SCI pathology, noting that it was often, but not exclusively linked with dpi. Evidence that the *Monocle3* trajectory analysis performed well[††] in this dataset comes from the trajectory following the differentiation of monocytes into either dendritic cells or macrophages. This differentiation is well described in the literature (Chomarat et al., 2000; Guilliams et al., 2014; Jakubzick et al., 2017), and was clearly observable in the dataset. Namely, the intermediate cluster, MC(2A), was largely contributed to by 3 dpi cells, the beginning of the first wave of macrophage infiltration (Alizadeh et al., 2019), and expressed GO terms related to cell differentiation. The two adjacent clusters on the UMAP, MC(3A) and DC(1), also clearly expressed macrophage and dendritic cell profiles, respectively. As such, this differentiation acted as a substitute for the ground truth and the *Monocle3* trajectory succeeded in predicting it. This supports the notion that the trajectory was biologically relevant in other partitions, such as the microglia clusters. Regardless, trajectory analysis remains only an estimate of the trajectory that a cell takes and should be interpreted with caution. Follow up experiments, such as BrdU labelling, could help validate a trajectory analysis.

To validate the key transcriptional patterns in this dataset, and to provide an element of spatial context, we utilised smFISH. While the smFISH data confirmed the general expression trends of the scRNAseq data (e.g., the increase in *Ms4a7*$^{+}$ infiltrating cells after SCI), the quantifications were not consistent. This could largely be attributed to the different anatomical regions sampled by the two approaches: scRNAseq collected cells from 5 mm of tissue, while smFISH considered only the lesion epicentre. However, given the slight

[††] Here “performed well” implies that the algorithm prediction biologically relevant transitions

discrepancies between resident myeloid cell *Serpine2* expression in the Ctrl tissue, there were likely additional confounding factors.

The scRNAseq and smFISH results may differ due to the technical biases inherent to each technique. For example, if a gene is expressed at levels too low to be reliably detected by the scRNAseq protocol, cells will be inadvertently labelled as negative for that gene (Kharchenko et al., 2014). smFISH has a higher sensitivity and much lower sparsity (Raj et al., 2008), which could explain why it found that nearly 100% of resident myeloid cells were *Serpine2*$^{+}$ in Ctrl samples, whereas the scRNAseq data reported only ~78%. scRNAseq, unlike smFISH, requires the isolation of single cells from the tissue, creating many opportunities for isolation-induced transcription, including the downregulation of homeostatic markers, or a selection-bias. For example, by 21 dpi, the spinal cord parenchyma is densely packed with newly proliferated and migrated cells, contributing to a glial scar and fibrotic core (Alizadeh et al., 2019). Thus, compared to Ctrl or acute phase tissue, cells from the lesion epicentre may be less easily extracted and less likely to remain intact. This theory is also supported by the low number of 21 dpi infiltrating myeloid cells recovered through scRNAseq, despite Veronica's quantification and the clear sense from the microscopy that the number of YFP^{+}/tdTomato^{-} cells is drastically increased at this stage (Table 4.1; Figure 4.6). The isolation process can also induce cell activation and thus transcription, confounding the scRNAseq results as discussed (3.2).

smFISH is not without its own caveats. The accurate partitioning of cells is an area of active research. In our case, Katherine was able to identify the cytoplasm of single cells using RFP/YFP intensity thresholding around single nuclei, however even with this fluorescent tool, the chronic, scarred tissue can provide segmentation challenges, especially between densely packed myeloid cells of the same fluorescent label. The smFISH protocol utilised in this study was only able to multiplex 4 mRNA probes, and, if combined with proteins, could detect just 5 targets in total. Given the modest modularity between many of the clusters, this drastically limited the ability to identify and confidently verify the presence of these scRNAseq-identified clusters *in situ.* Spatial techniques with greater multiplexing capabilities do exist, for e.g. imaging mass cytometry, which can detect up to 32 protein targets (Giesen et al., 2014), or GeoMX WTA, which can effectively sequence the entire transcriptome *in situ* (Roberts et al.,

2021). However, with that degree of multiplexing, such an approach would simply perform better than scRNAseq at characterising the heterogeneity in myeloid cells after SCI, particularly given the importance of spatial context in this pathology.

A major focus of this chapter was the dynamic expression of *Fabp5* after SCI. In both resident and peripheral myeloid cells, *Fabp5* correlated with bifurcations in the trajectory analysis. For microglia, *Fabp5* expression was linked to a shift in the trajectory, away from the neuroprotective MG(3A), and towards the cytotoxic, MGnD-like, foam cell-like MG(3B) phenotype. In monocyte-derived cells, *Fabp5* expression marked the fate of differentiation into macrophages with foam cell-like profiles, as opposed to pro-regenerative dendritic cells (Yaguchi et al., 2009). Under both homeostatic and activated conditions, peripheral myeloid cells express *Fabp5* (Moore et al., 2015), however microglia express *Fabp5* only during development or upon activation (Hammond et al., 2019), suggesting a specific role for *Fabp5* in activated myeloid cells after SCI, which has yet to be discerned (Peruzzotti-Jametti et al., 2021).

FABPs are a family of cytosolic lipid chaperones that reversibly bind to hydrophobic molecules, such as free FAs. FABPs transport their cargo to specific nuclear compartments, such as the endoplasmic reticulum for lipoprotein, triglyceride, and cholesterol synthesis (Furuhashi & Hotamisligil, 2008; Peruzzotti-Jametti et al., 2021; Wu et al., 2010), or to nuclear receptors such as PPARγ (Levi et al., 2013). As such, FABP5 loss of function has been shown to lead to a build-up of intracellular FAs and a decrease in cholesterol (Zhang et al., 2004). In bone marrow-derived macrophages (BMDM) with FABP5 loss of function, inflammatory (LPS and IFN-γ) or anti-inflammatory (IL-4) stimulation results in significantly higher expression of anti-inflammatory factors (Moore et al., 2015; Y. Zhang et al., 2014). These findings demonstrate that loss of FABP5 promotes anti-inflammatory responses in macrophages. Thus, while little is known about the role of FABP5 in microglia, it represents an interesting target that could be manipulated to alter lipid metabolism and reduce chronic neuroinflammation.

In regulatory T-cells *in vitro*, FABP5 loss of function results in decreased OXPHOS and increased glycolysis, altered mitochondrial morphology, and impaired lipid metabolism (Field

et al., 2020). The same study found that a decrease in lipid availability both *in vitro* and in the tumour microenvironment resulted in an upregulation of *Fabp5* in regulatory T-cells. Another study in carcinoma cells found that the transcription factor, NF-κB, induces *Fabp5* expression (Kannan-Thulasiraman et al., 2010). Either mechanism may be responsible for the upregulation of *Fabp5* observed in clusters MG(3B) and MC(3A). However, given that microglia and macrophages play a central role in clearing lipid-based myelin debris after SCI (Kopper & Gensel, 2018), and phagocytosed myelin has been demonstrated to upregulated *Msr1* and consequently NF-κB activity in BMDM *in vitro* (Kong et al., 2020), the latter is the more likely candidate. This notion is further supported by GO terms for the positive regulation of NF-κB signalling in MG(1) and MC(2A), two clusters that precede the *Fabp5*^{hi} clusters in the trajectory analysis and in dpi (Figure 5.16C; Figure 5.30A).

Msr1 was recently demonstrated to promote the formation of foam cells and neuronal apoptosis through myelin-mediated NF-κB signalling in weight-drop mouse models of T10 contusive SCI (Kong et al., 2020). The global KO of *Msr1* improved functional outcomes. The authors focused their mechanistic studies on *in vitro* BMDM, but I observed evidence of this signalling pathway in both microglia and macrophages. Similar to the observations of Kong et al., the foam cell-like clusters in this SCI myeloid cell atlas appeared to be cytotoxic and pro-inflammatory, and this phenotype persisted into the chronic phase at 21 dpi. To know whether this phenotype is naturally resolved during the course of the chronic disease would require longer term investigations.

As discussed, the microglia foam cell-like cluster, MG(3B), also adopted the previously described MGnD (Krasemann et al., 2017) and to a lesser extent DAM2 (Keren-Shaul et al., 2017) phenotypes. While at a first glance this may seem inconsistent, upon closer reflection, many of the marker genes are repeated across the phenotypes (*Clec7a, Trem2, Spp1, Lgals3, Cd9, Itgax*) and are involved in lipid metabolism. Interestingly, the LDAM (Marschallinger et al., 2020) phenotype was not observed in this dataset, suggesting that the lipid droplet accumulation in aged mice is different than the myelin debris-induced foam cell-like profiles observed in the myeloid cell atlas. In support of this, Marschallinger et al., did not observe substantial overlap between their LDAM signature and the MGnD and DAM phenotypes in their dataset either.

Together this data suggests that *Fabp5* may serve as a marker of foam cell-like phenotypes in the CNS, which is a distinct phenotype from age-related LDAM but overlaps with the AD MGnD microglia. Future studies could investigate the mechanism behind the bifurcation in the MG(3) trajectory – what allows certain cells to forgo the MG(3B) foam cell-like state and adopt the neuroprotective MG(3A) phenotype? Given that *Fabp5* upregulation precedes the foam cell state, myeloid cell *Fabp5* loss of function could be investigated for its potential to prevent foam cell formation, cytotoxicity reduction, and ultimately improved functional outcomes.

Chapter 6 DISCUSSION

“I didn’t want to just know names of things. I remember really wanting to know how it all worked.”

— Elizabeth Blackburn

scRNAseq is a powerful descriptive tool that can help deepen our understanding of cellular level processes in health and disease. In the case of SCI and this myeloid cell atlas, scRNAseq has provided a new level of understanding by piecing together the findings of decades of research and summarising them within a single dataset. This powerful, high-throughput technology now routinely builds such platforms upon which we can formulate and test novel biological questions, ever strengthening our understanding of cellular biology. Naturally, the approach is not without limitations, but the rapidly evolving field of single-cell data science provides exciting opportunities for future omics studies in SCI.

6.1 Caveats of scRNAseq in the Context of the SCI Myeloid Cell Atlas

As a methodology, scRNAseq has many shortcomings, some of which are attributable to the protocol, while others are more fundamental. In addition to these broader limitations, scRNAseq applied to SCI has its own unique challenges, which were reflected in this time-resolved SCI myeloid cell atlas.

The single cell isolation protocol inherent to scRNAseq presents several challenges that can confound the downstream interpretation. As discussed throughout this thesis, extracting single cells from tissue inevitably damages the cells and, particularly in the case of microglia, activates them (Haimon et al., 2018; Hammond et al., 2019; Marsh et al., 2020; Milich et al., 2021). Certain steps can be taken to minimise this activation, such as performing the protocol at 4°C, but this mitigation strategy comes with other drawbacks. For example, cold dissociation has been demonstrated to reduce the efficiency of extracting densely packed tissues, such as that observed in the subacute and chronic SCI phase (Denisenko et al., 2020). This can produce a sampling bias even within the same cell type where certain cell states – based on morphology, surrounding tissue composition, or other unknown factors – have different probabilities of surviving the isolation protocol and giving rise to high quality cell libraries (Denisenko et al., 2020; Haimon et al., 2018). In this myeloid cell atlas, this sampling bias might account for the diminished 21 dpi samples, or the inconsistent proportions of foam cell-like MG(3B) cells observed in the scRNAseq *vs* smFISH data.

The mechanical homogenisation step required for cold-dissociation shears many fragile CNS cells, such as neurons and astrocytes (Marsh et al., 2020), generating substantial ambient mRNA as observed in *Chapter 3* and *4.* This background undoubtably changes with the tissue composition, and so in SCI, dpi, which could confound DEG analysis or enhance condition-specific shifts in the data. Algorithms, such as SoupX (Young & Behjati, 2018), have been designed to estimate and subtract this background. In a previous analysis workflow (Hamel et al., 2020), I applied this algorithm in the SCI myeloid cell atlas, but the general conclusions were consistent with this dissertation, suggesting they were not driven by ambient mRNA.

After the completion of the sample collection phase of this study, a new protocol was developed to mitigate the biases arising from 4°C mechanical homogenisation. This approach utilises transcription and translation inhibitors to prevent *ex vivo*-induced gene products during enzymatic dissociation (Marsh et al., 2020). As previously discussed, such a protocol would have reduced *ex vivo*-induced transcription, which manifests itself in the study as the cycling in the Ctrl sample, G11.D2, and the increased *Apoe* expression compared to the smFISH data. However, these technical artefacts appeared to be minimal compared to other reports (Figure 3.2) (Milich et al., 2021). Critically, the Marsh et al., protocol would have likely improved the quality of the isolated myeloid cells and reduced the volume of ambient mRNA, ultimately resulting in a cleaner, larger myeloid cell atlas.

After single cells have been isolated from the tissue, the protocols for droplet encapsulation, barcoding, and the sequencing itself all present opportunities for complication. Most relevant to this SCI myeloid cell atlas is the sparsity of scRNAseq data: technically and biologically driven zero counts in the gene expression data previously referred to as “dropout events” (Hicks et al., 2017; Lähnemann et al., 2020). The technically-driven false negatives are correlated with gene expression level, with lowly expressed genes resulting in greater sparsity (Hicks et al., 2017). In this myeloid cell atlas, I noticed the absence of several expected genes, such as the microglial-specific transcription factor *Sall1*, or the conceivable expression of other MGnD genes like *Lilrb4a*. The degree of sparsity has been shown to vary from cell to cell and several algorithms have been designed to exploit this to impute the missing values, but with the relatively high risk of generating false-positive results that lack reproducibility (Andrews & Hemberg, 2019). This is, however, an active area of research (Lähnemann et al.,

2020). Sparsity can also be biologically driven through stochastic transcriptional bursting (Marinov et al., 2014), highlighting some broader issues surrounding transcriptional methods.

scRNAseq only provides us with a snapshot of a cell's state, with no context as to the cells' previous or future states. Live, repeatable sampling is on the horizon, but has certainly not yet reached a 10x Chromium scRNAseq level of accessibility and robustness (Chen et al., 2021). In the meantime, approaches such as trajectory analysis allow us to take these snapshots of thousands of cells, each in a slightly different state, and deduce the transcriptional changes that an individual cell might pass through. As discussed in *Chapter 5*, this is not without its caveats, but remains a useful tool.

Typically, and certainly in the case of this myeloid cell atlas, scRNAseq characterises cells on an mRNA-level to elucidate protein-level functions, in lieu of equivalently high-throughput protein measurements. This scRNAseq approach makes several key assumptions, including that each detected mRNA would have elicited a functional protein. Numerous studies have reported that scRNAseq quantifications do not mirror protein abundance and that this relationship is largely context-dependent, fluctuating with the availability of resources for protein synthesis (Y. Liu et al., 2016; Reimegård et al., 2021). Further deepening the disconnect, mRNA is much less stable than protein, and its half-life varies by gene (Schwanhäusser et al., 2011). At least for transcription factors, protein levels more accurately reflect their downstream effects (Reimegård et al., 2021) and many key cellular-level processes, including cell cycling, apoptosis, and innate inflammatory responses are mediated not merely by the presence of a protein but by its post-translational modifications (Burke et al., 2012; J. Liu et al., 2016; Vucic et al., 2011). Thus, approaches that combine single cell transcriptomics with highly multiplexed protein-level analysis will likely become the standard in the field. Currently, a key barrier to the widespread use of this approach is the restriction of protein detection to either cell surface proteins, through techniques including CITE-seq (Stoeckius et al., 2017) and REAP-seq (Peterson et al., 2017), or intracellular proteins using approaches such as SPARC (Reimegård et al., 2021). Notably, *in situ* methods to combine transcriptomics with multiplexed protein-level analysis, such as DBiT-seq (Liu et al., 2020), also represent an area of active research.

Another assumption made during the analysis of scRNAseq data is that the mRNA within a cell reflects the cell's own transcriptional activity. Pertinently for the characterisation of phagocytes, this is not necessarily true. A recent study compared the transcriptome and RiboTag-acquired (Sanz et al., 2009) translatome of homeostatic and LPS-treated microglia and found significant phagocytic cargo contamination in both conditions (Haimon et al., 2018). Currently no strategies exist to reliably identify phagocytic cargo in scRNAseq data, one can merely consider this caveat during computational analysis and aim to perform mRNA and protein-level validation of key findings.

Other limitations of scRNAseq include the descriptive nature and loss of spatial context, which have been extensively discussed throughout the dissertation. Understanding that gene products represent only one aspect of cellular biology is also critical and has called for the establishment of multi-omics approaches. Such techniques aim to examine the transcriptome alongside other big datasets, such as the epigenome or metabolome, and represent a relevant and exciting field within single cell biology (Luecken & Theis, 2019).

6.2 Future Directions of Transcriptomics in SCI

This study employed scRNAseq to investigate myeloid cells at an unprecedented resolution in the subacute and chronic phase of SCI. While this transcriptomics dataset represents a valuable resource for future investigations, advancements in the field, particularly multi-omics and *in situ* technologies present an opportunity to further enhance our understanding of the SCI pathology.

Chronic inflammation and activation are hallmarks of the SCI pathology (Alizadeh et al., 2019). Profiling the epigenetic changes that give rise to these chronic cell states, alongside their transcriptional characterisations, might enable us to target dysregulated pathways before they result in destructive phenotypes. For example, understanding the epigenetic changes that precede the trajectory-predicted split of MG(3) cells towards the neuroprotective or neurodegenerative foam cell-like state could allow us to manipulate the epigenome and enrich the spinal cord tissue for the neuroprotective cell state. However, as discussed, spatial context is critical to fully understanding the SCI pathology and single cell epigenomic technologies, such as scATAC-seq (Buenrostro et al., 2015), or the 10X Chromium Single Cell

Multiome ATAC + Gene Expression, would relinquish the spatial information, just as scRNAseq does. Furthermore, these approaches would face the same isolation challenges as this current study: low cell yields, sampling bias, and isolation-induced transcription. Thus, future omics studies of the SCI pathology should endeavour to collect *in situ* data.

Currently *in situ* techniques are restricted to a single modality (except DBiT-seq), but nonetheless *in situ* methods, like CUT&RUN (Skene & Henikoff, 2017) for *in situ* chromatin mapping or GeoMX WTA (Roberts et al., 2021) for *in situ* transcriptomics, would add valuable insight to the SCI pathology. Both *in situ* approaches would allow for the characterisation of all SCI cell types, in parallel, allowing for the investigation of multiple cell types and their interactions. The significance of profiling all CNS cell types after SCI was demonstrated by Milich et al., 2021, who gained valuable insight through ligand-receptor analysis on their scRNAseq data. Further adding to the importance of *in situ* techniques, GeoMX WTA is capable of profiling the transcriptome in FFPE and fresh frozen tissue, which has made transcriptomics studies of post-mortem human tissue feasible. This is particularly important as, while animal models of SCI are critical for mechanistic studies that are not possible in humans, relying on these models for descriptive studies is suboptimal, and the translatability of transcriptomics studies between mice and humans is a matter of debate (Osterburg et al., 2013; Seok et al., 2013).

Omics approaches are descriptive by nature, proficient at generating hypotheses but unable to provide definitive causal evidence. Performing these technologies under controlled treatment conditions would allow for greater mechanistic discoveries. Given that studies of the SCI pathology are currently restricted to animal models, high-throughput perturbation approaches are not realistic. Thus, the treatments must be chosen with great care. The growing body of high-throughput data is quickly expanding and publicly available databases such as the GEO and the Open Data Commons for Spinal Cord Injury (https://odc-sci.org/) serve as valuable resources for hypothesis generation and the identification of potential treatment candidates. With this growing foundation of high-throughput data, researchers and journals should begin to consider the futility of generating omics data just for the sake of high impact factor publications.

6.3 THESIS SUMMARY

In this thesis, I utilised a well-established mouse model of T12 contusion SCI on two different mouse strains to isolate and scRNAseq myeloid cells across the SCI pathology, ultimately generating a time-resolved SCI myeloid cell atlas. After careful pre-processing and analysis, this atlas replicated many previously reported myeloid cell characteristics, both in the context of SCI and other CNS pathologies, summarising decades of research within a single dataset. Furthermore, the atlas provided an unprecedented level of resolution in the subacute and chronic phase and suggested that the previously described macrophage foam cell state is also adopted by microglia. These foam cell states appear to be triggered by the phagocytosis of myelin debris in the acute phase, which upregulates *Msr1* and subsequently induces NF-κB signalling. This may ultimately result in a lipid-processing foam cell-like microglia state in the subacute and chronic phase. These cells also appear to adopt a neurodegenerative MGnD-like profile, suggesting a detrimental role in the SCI response. Furthermore, lipid-laden macrophages have been described to persist for at least one year in human SCI, implicating them in the chronic neuroinflammatory response (Fleming et al., 2006). In the mouse SCI myeloid cell atlas, these foam-like cells can be identified by their *Fabp5* expression, but the exact role of this gene in the context of SCI requires further investigation.

The contusion model of SCI mimics the most common form of human SCI and, thanks to the computerised impactor, is highly reproducible. In this dissertation I demonstrated this reproducibility on both a behavioural and single-cell level. Animal models of SCI are a valuable resource to study the SCI pathology and do not yet, unfortunately, have a suitable replacement. Continuous refinement of the model serves to improve its value in studying the pathology, while minimising animal suffering. In this study, I worked to add to that refinement and endeavoured to make the high-throughput data easily accessible, to reduce the need for future mouse models of SCI.

I first applied this mouse model of SCI to a *Cx3cr1* mouse, allowing for the efficient isolation and scRNAseq of *Cx3cr1*$^{+}$ myeloid cells after SCI. I collected cells from laminectomy-only controls, as well as from the acute, subacute, and early chronic phases of SCI, taking significant care to minimise extraction-induced transcription in this notoriously responsive cell populations. Using this *Cx3cr1* data, I established a pre-processing, QC, and preliminary

analysis workflows based on published methods, most of which have recently been summarised in an online book (Amezquita et al., 2020). Integration of the batches comprising this dataset proved challenging, emphasising the importance of thoughtful experimental design in scRNAseq studies. Despite this caveat, this dataset contained nearly 12,000 high quality cells and exhibited the expected cell types, including both infiltrating and resident myeloid cells. From this dataset alone, the macrophage *vs* microglia identity was, in certain cases, ambiguous, inciting the upgrade to a fate-mapping mouse line.

I repeated the protocol and workflows established for the *Cx3cr1* mouse on a transgenic fate mapping mouse line. This line differentially labelled infiltrating *vs* resident myeloid cells, which we validated via smFISH and IF, allowing for the unambiguous annotation of these cell types both *ex vivo* and *in silico.* The MNN integration of the *Cx3cr1* and *Cremato* datasets was made easier by the fact that both datasets comprised the same SCI time points, allowing each dpi to be appropriately merged with its counterpart. I investigated the variance lost within each strain during correction and found that ≤ 2% of variance was removed, well within the recommended upper limit of 10% (Amezquita et al., 2020). Furthermore, the MNN correction correctly merged cell types and dpi across the two strains, suggesting that the variance lost was due to strain or batch effect. I made this SCI myeloid cell atlas publicly available on the GEO (GSE159638) and a Shiny-based interactive web app (https://marionilab.cruk.cam.ac.uk/SCI_Myeloid_Cell_Atlas/), allowing other researchers to benefit from this study.

After compiling the SCI myeloid cell atlas, I confirmed the presence of known phenotypes, including the activation-induced downregulation of homeostatic microglia markers, the adoption of DAM-like phenotypes, the lack of canonical M1/M2 polarisation, and acute phase proliferation. Then I performed unsupervised clustering and trajectory analysis and characterised each cluster, ultimately generating a map of myeloid cell states across the SCI pathology. We verified several of the expression-based findings using smFISH and IF but ultimately further investigations *in vitro* and *in vivo* are required to validate any functional conclusions.

REFERENCES

Abels, E. R., Nieland, L., Hickman, S., Broekman, M. L. D., El Khoury, J., & Maas, S. L. N. (2021). Comparative Analysis Identifies Similarities between the Human and Murine Microglial Sensomes. *International Journal of Molecular Sciences, 22*(3), 1495. https://doi.org/10.3390/ijms22031495

Ahmed, N., & Weidemann, M. J. (1994, 1994/06/01/). Purine metabolism in promyelocytic HL60 and dimethylsulphoxide-differentiated HL60 cells. *Leukemia Research, 18*(6), 441-451. https://doi.org/https://doi.org/10.1016/0145-2126(94)90080-9

Ahuja, C. S., Nori, S., Tetreault, L., Wilson, J., Kwon, B., Harrop, J., Choi, D., & Fehlings, M. G. (2017). Traumatic Spinal Cord Injury—Repair and Regeneration. *Neurosurgery, 80*(3S), S9-S22. https://doi.org/10.1093/neuros/nyw080

Akhtar, A. Z., Pippin, J. J., & Sandusky, C. B. (2008). Animal models in spinal cord injury: a review. *Rev Neurosci, 19*(1), 47-60. https://doi.org/10.1515/revneuro.2008.19.1.47

Alizadeh, A., Dyck, S. M., & Karimi-Abdolrezaee, S. (2019, 2019-March-22). Traumatic Spinal Cord Injury: An Overview of Pathophysiology, Models and Acute Injury Mechanisms [Review]. *Frontiers in Neurology, 10*(282). https://doi.org/10.3389/fneur.2019.00282

Amezquita, R. A., Lun, A. T. L., Becht, E., Carey, V. J., Carpp, L. N., Geistlinger, L., Marini, F., Rue-Albrecht, K., Risso, D., Soneson, C., Waldron, L., Pagès, H., Smith, M. L., Huber, W., Morgan, M., Gottardo, R., & Hicks, S. C. (2020, 2020/02/01). Orchestrating single-cell analysis with Bioconductor. *Nature Methods, 17*(2), 137-145. https://doi.org/10.1038/s41592-019-0654-x

Amici, S. A., Dong, J., & Guerau-de-Arellano, M. (2017, 2017-November-10). Molecular Mechanisms Modulating the Phenotype of Macrophages and Microglia [Review]. *Frontiers in Immunology, 8*(1520). https://doi.org/10.3389/fimmu.2017.01520

Anderson, M. A., Burda, J. E., Ren, Y., Ao, Y., O'Shea, T. M., Kawaguchi, R., Coppola, G., Khakh, B. S., Deming, T. J., & Sofroniew, M. V. (2016, Apr 14). Astrocyte scar formation aids central nervous system axon regeneration. *Nature, 532*(7598), 195-200. https://doi.org/10.1038/nature17623

Andrews, T., & Hemberg, M. (2019). False signals induced by single-cell imputation [version 2; peer review: 4 approved]. *F1000Research, 7*(1740). https://doi.org/10.12688/f1000research.16613.2

Arneson, D., Zhang, G., Ying, Z., Zhuang, Y., Byun, H. R., Ahn, I. S., Gomez-Pinilla, F., & Yang, X. (2018). Single cell molecular alterations reveal target cells and pathways of concussive

brain injury. *Nat Commun, 9*(1), 3894-3894. https://doi.org/10.1038/s41467-018-06222-0

Atagi, Y., Liu, C.-C., Painter, M. M., Chen, X.-F., Verbeeck, C., Zheng, H., Li, X., Rademakers, R., Kang, S. S., Xu, H., Younkin, S., Das, P., Fryer, J. D., & Bu, G. (2015, 2015/10/23/). Apolipoprotein E Is a Ligand for Triggering Receptor Expressed on Myeloid Cells 2 (TREM2)*. *Journal of Biological Chemistry, 290*(43), 26043-26050. https://doi.org/https://doi.org/10.1074/jbc.M115.679043

Au - Richner, M., Au - Jager, S. B., Au - Siupka, P., & Au - Vaegter, C. B. (2017, 2017/01/22/). Hydraulic Extrusion of the Spinal Cord and Isolation of Dorsal Root Ganglia in Rodents. *JoVE*(119), e55226. https://doi.org/doi:10.3791/55226

Bais, A. S., & Kostka, D. (2020). scds: computational annotation of doublets in single-cell RNA sequencing data. *Bioinformatics (Oxford, England), 36*(4), 1150-1158. https://doi.org/10.1093/bioinformatics/btz698

Basso, M. D., Fisher, L. C., Anderson, A. J., Jakeman, L. B., Mctigue, D. M., & Popovich, P. G. (2006). Basso Mouse Scale for Locomotion Detects Differences in Recovery after Spinal Cord Injury in Five Common Mouse Strains. *Journal of neurotrauma, 23*(5), 635-659. https://doi.org/10.1089/neu.2006.23.635

Beck, K. D., Nguyen, H. X., Galvan, M. D., Salazar, D. L., Woodruff, T. M., & Anderson, A. J. (2010, Feb). Quantitative analysis of cellular inflammation after traumatic spinal cord injury: evidence for a multiphasic inflammatory response in the acute to chronic environment. *Brain, 133*(Pt 2), 433-447. https://doi.org/10.1093/brain/awp322

Behjati, S., & Frank, M. H. (2009). The effects of tamoxifen on immunity. *Current medicinal chemistry, 16*(24), 3076-3080. https://doi.org/10.2174/092986709788803042

Bellver-Landete, V., Bretheau, F., Mailhot, B., Vallières, N., Lessard, M., Janelle, M.-E., Vernoux, N., Tremblay, M.-È., Fuehrmann, T., Shoichet, M. S., & Lacroix, S. (2019, 2019/01/31). Microglia are an essential component of the neuroprotective scar that forms after spinal cord injury. *Nat Commun, 10*(1), 518. https://doi.org/10.1038/s41467-019-08446-0

Bennett, F. C., Bennett, M. L., Yaqoob, F., Mulinyawe, S. B., Grant, G. A., Hayden Gephart, M., Plowey, E. D., & Barres, B. A. (2018, Jun 27). A Combination of Ontogeny and CNS Environment Establishes Microglial Identity. *Neuron, 98*(6), 1170-1183.e1178. https://doi.org/10.1016/j.neuron.2018.05.014

Bernier, L.-P., York, E. M., Kamyabi, A., Choi, H. B., Weilinger, N. L., & MacVicar, B. A. (2020, 2020/03/25). Microglial metabolic flexibility supports immune surveillance of the brain parenchyma. *Nat Commun, 11*(1), 1559. https://doi.org/10.1038/s41467-020-15267-z

Biyani, A., & Masry, W. S. E. (1994, 1994/11/01). Post-traumatic syringomyelia: a review of the literature. *Spinal Cord, 32*(11), 723-731. https://doi.org/10.1038/sc.1994.117

Boivin, A., Pineau, I., Barrette, B., Filali, M., Vallières, N., Rivest, S., & Lacroix, S. (2007). Toll-Like Receptor Signaling Is Critical for Wallerian Degeneration and Functional Recovery after Peripheral Nerve Injury. *The Journal of Neuroscience, 27*(46), 12565. https://doi.org/10.1523/JNEUROSCI.3027-07.2007

Bowes, A. L., & Yip, P. K. (2014, Nov 1). Modulating inflammatory cell responses to spinal cord injury: all in good time. *Journal of neurotrauma, 31*(21), 1753-1766. https://doi.org/10.1089/neu.2014.3429

Bradbury, E. J., & Burnside, E. R. (2019, 2019/08/28). Moving beyond the glial scar for spinal cord repair. *Nat Commun, 10*(1), 3879. https://doi.org/10.1038/s41467-019-11707-7

Braga, A., Bandiera, S., Verheyen, J., Hamel, R., Rutigliani, C., Edenhofer, F., Smith, J. A., & Pluchino, S. (2020, 2020/08/05/). Combination of in situ Lcn2 pRNA-RNAi nanotherapeutics and iNSC transplantation ameliorates experimental spinal cord injury in mice. *Molecular Therapy*. https://doi.org/https://doi.org/10.1016/j.ymthe.2020.08.001

Brennan, F. H., & Popovich, P. G. (2018, Jun). Emerging targets for reprograming the immune response to promote repair and recovery of function after spinal cord injury. *Curr Opin Neurol, 31*(3), 334-344. https://doi.org/10.1097/wco.0000000000000550

Bruttger, J., Karram, K., Wörtge, S., Regen, T., Marini, F., Hoppmann, N., Klein, M., Blank, T., Yona, S., Wolf, Y., Mack, M., Pinteaux, E., Müller, W., Zipp, F., Binder, H., Bopp, T., Prinz, M., Jung, S., & Waisman, A. (2015, 2015/07/21/). Genetic Cell Ablation Reveals Clusters of Local Self-Renewing Microglia in the Mammalian Central Nervous System. *Immunity, 43*(1), 92-106. https://doi.org/https://doi.org/10.1016/j.immuni.2015.06.012

Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., Chang, H. Y., & Greenleaf, W. J. (2015, Jul 23). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature, 523*(7561), 486-490. https://doi.org/10.1038/nature14590

Burda, J. E., & Sofroniew, M. V. (2014, Jan 22). Reactive gliosis and the multicellular response to CNS damage and disease. *Neuron, 81*(2), 229-248. https://doi.org/10.1016/j.neuron.2013.12.034

Burke, J. R., Hura, G. L., & Rubin, S. M. (2012, Jun 1). Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control. *Genes Dev, 26*(11), 1156-1166. https://doi.org/10.1101/gad.189837.112

Buss, A., Pech, K., Kakulas, B. A., Martin, D., Schoenen, J., Noth, J., & Brook, G. A. (2009, 2009/07/15). NG2 and phosphacan are present in the astroglial scar after human traumatic spinal cord injury. *BMC Neurology, 9*(1), 32. https://doi.org/10.1186/1471-2377-9-32

Butovsky, O., Jedrychowski, M. P., Moore, C. S., Cialic, R., Lanser, A. J., Gabriely, G., Koeglsperger, T., Dake, B., Wu, P. M., Doykan, C. E., Fanek, Z., Liu, L., Chen, Z., Rothstein, J. D., Ransohoff, R. M., Gygi, S. P., Antel, J. P., & Weiner, H. L. (2014, 2014/01/01). Identification of a unique TGF-β–dependent molecular and functional signature in microglia. *Nature Neuroscience, 17*(1), 131-143. https://doi.org/10.1038/nn.3599

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019, 2019/02/01). The single-cell transcriptional landscape of mammalian organogenesis. *Nature, 566*(7745), 496-502. https://doi.org/10.1038/s41586-019-0969-x

Carpenter, R. S., Jiang, R. R., Brennan, F. H., Hall, J. C. E., Gottipati, M. K., Niewiesk, S., & Popovich, P. G. (2019, 2019/12/13). Human immune cells infiltrate the spinal cord and impair recovery after spinal cord injury in humanized mice. *Scientific Reports, 9*(1), 19105. https://doi.org/10.1038/s41598-019-55729-z

Carpenter, R. S., Kigerl, K. A., Marbourg, J. M., Gaudet, A. D., Huey, D., Niewiesk, S., & Popovich, P. G. (2015). Traumatic spinal cord injury in mice with human immune systems. *Experimental Neurology, 271*, 432-444. https://doi.org/10.1016/j.expneurol.2015.07.011

Cekanaviciute, E., & Buckwalter, M. S. (2016, 2016/10/01). Astrocytes: Integrative Regulators of Neuroinflammation in Stroke and Other Neurological Diseases. *Neurotherapeutics, 13*(4), 685-701. https://doi.org/10.1007/s13311-016-0477-8

Chandak, P. G., Radovic, B., Aflaki, E., Kolb, D., Buchebner, M., Fröhlich, E., Magnes, C., Sinner, F., Haemmerle, G., Zechner, R., Tabas, I., Levak-Frank, S., & Kratky, D. (2010, Jun 25). Efficient phagocytosis requires triacylglycerol hydrolysis by adipose triglyceride lipase. *The Journal of biological chemistry, 285*(26), 20192-20201. https://doi.org/10.1074/jbc.M110.107854

Chen, K., Deng, S., Lu, H., Zheng, Y., Yang, G., Kim, D., Cao, Q., & Wu, J. Q. (2013, 08/09 03/14/received
07/13/accepted). RNA-Seq Characterization of Spinal Cord Injury Transcriptome in Acute/Subacute Phases: A Resource for Understanding the Pathology at the Systems Level. *PloS one, 8*(8), e72567. https://doi.org/10.1371/journal.pone.0072567

Chen, W., Guillaume-Gentil, O., Dainese, R., Rainer, P. Y., Zachara, M., Gäbelein, C. G., Vorholt, J. A., & Deplancke, B. (2021). Genome-wide molecular recording using Live-seq. *bioRxiv*, 2021.2003.2024.436752. https://doi.org/10.1101/2021.03.24.436752

Cheriyan, T., Ryan, D. J., Weinreb, J. H., Cheriyan, J., Paul, J. C., Lafage, V., Kirsch, T., & Errico, T. J. (2014, Aug). Spinal cord injury models: a review. *Spinal Cord, 52*(8), 588-595. https://doi.org/10.1038/sc.2014.91

Chomarat, P., Banchereau, J., Davoust, J., & Karolina Palucka, A. (2000, 2000/12/01). IL-6 switches the differentiation of monocytes from dendritic cells to macrophages. *Nature Immunology, 1*(6), 510-514. https://doi.org/10.1038/82763

Choo, A. M.-T., Liu, J., Liu, Z., Dvorak, M., Tetzlaff, W., & Oxland, T. R. (2009, 2009/06/30/). Modeling spinal cord contusion, dislocation, and distraction: Characterization of vertebral clamps, injury severities, and node of Ranvier deformations. *Journal of Neuroscience Methods, 181*(1), 6-17. https://doi.org/https://doi.org/10.1016/j.jneumeth.2009.04.007

Colombo, E., & Farina, C. (2016, 2016/09/01/). Astrocytes: Key Regulators of Neuroinflammation. *Trends Immunol, 37*(9), 608-620. https://doi.org/https://doi.org/10.1016/j.it.2016.06.006

Courtine, G., Bunge, M. B., Fawcett, J. W., Grossman, R. G., Kaas, J. H., Lemon, R., Maier, I., Martin, J., Nudo, R. J., Ramon-Cueto, A., Rouiller, E. M., Schnell, L., Wannier, T., Schwab, M. E., & Edgerton, V. R. (2007). Can experiments in nonhuman primates expedite the translation of treatments for spinal cord injury in humans? *Nature Medicine, 13*(5), 561-566. https://doi.org/10.1038/nm1595

Cregg, J. M., DePaul, M. A., Filous, A. R., Lang, B. T., Tran, A., & Silver, J. (2014, 2014/03/01/). Functional regeneration beyond the glial scar. *Experimental Neurology, 253*, 197-207. https://doi.org/https://doi.org/10.1016/j.expneurol.2013.12.024

Cusimano, M., Biziato, D., Brambilla, E., Donegà, M., Alfaro-Cervello, C., Snider, S., Salani, G., Pucci, F., Comi, G., Garcia-Verdugo, J. M., De Palma, M., Martino, G., & Pluchino, S. (2012). Transplanted neural stem/precursor cells instruct phagocytes and reduce secondary tissue damage in the injured spinal cord. *Brain, 135*(2), 447-460. https://doi.org/10.1093/brain/awr339

David, S., & Kroner, A. (2011, 2011/07/01). Repertoire of microglial and macrophage responses after spinal cord injury. *Nature Reviews Neuroscience, 12*(7), 388-399. https://doi.org/10.1038/nrn3053

Deczkowska, A., Keren-Shaul, H., Weiner, A., Colonna, M., Schwartz, M., & Amit, I. (2018, 2018/05/17/). Disease-Associated Microglia: A Universal Immune Sensor of Neurodegeneration. *Cell, 173*(5), 1073-1081. https://doi.org/https://doi.org/10.1016/j.cell.2018.05.003

Denisenko, E., Guo, B. B., Jones, M., Hou, R., de Kock, L., Lassmann, T., Poppe, D., Clément, O., Simmons, R. K., Lister, R., & Forrest, A. R. R. (2020, 2020/06/02). Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus

RNA-seq workflows. *Genome Biology, 21*(1), 130. https://doi.org/10.1186/s13059-020-02048-6

Devanney, N. A., Stewart, A. N., & Gensel, J. C. (2020). Microglia and macrophage metabolism in CNS injury and disease: The role of immunometabolism in neurodegeneration and neurotrauma. *Experimental Neurology, 329*, 113310-113310. https://doi.org/10.1016/j.expneurol.2020.113310

DeVivo, M. J. (2012, 01/24/online). Epidemiology of traumatic spinal cord injury: trends and future implications [Review]. *Spinal Cord, 50*, 365. https://doi.org/10.1038/sc.2011.178

Donnelly, D. J., & Popovich, P. G. (2008, Feb). Inflammation and its role in neuroprotection, axonal regeneration and functional recovery after spinal cord injury. *Exp Neurol, 209*(2), 378-388. https://doi.org/10.1016/j.expneurol.2007.06.009

Dubbelaar, M. L., Kracht, L., Eggen, B. J. L., & Boddeke, E. W. G. M. (2018, 2018-July-31). The Kaleidoscope of Microglial Phenotypes [Review]. *Frontiers in Immunology, 9*(1753). https://doi.org/10.3389/fimmu.2018.01753

Dunham, K. A., Siriphorn, A., Chompoopong, S., & Floyd, C. L. (2010, 2010/11/01). Characterization of a Graded Cervical Hemicontusion Spinal Cord Injury Model in Adult Male Rats. *Journal of neurotrauma, 27*(11), 2091-2106. https://doi.org/10.1089/neu.2010.1424

Durruthy, J. (2019). *Getting Started With Single Cell Expression*. https://www.10xgenomics.com/products/single-cell-gene-expression

Dworzak, J., Renvoisé, B., Habchi, J., Yates, E. V., Combadière, C., Knowles, T. P., Dobson, C. M., Blackstone, C., Paulsen, O., & Murphy, P. M. (2015). Neuronal Cx3cr1 Deficiency Protects against Amyloid β-Induced Neurotoxicity. *PloS one, 10*(6), e0127730. https://doi.org/10.1371/journal.pone.0127730

E Hirbec, H., Noristani, H. N., & Perrin, F. E. (2017). Microglia Responses in Acute and Chronic Neurological Diseases: What Microglia-Specific Transcriptomic Studies Taught (and did Not Teach) Us. *Frontiers in aging neuroscience, 9*, 227-227. https://doi.org/10.3389/fnagi.2017.00227

Efremova, M., Vento-Tormo, M., Teichmann, S. A., & Vento-Tormo, R. (2020, 2020/04/01). CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nature Protocols, 15*(4), 1484-1506. https://doi.org/10.1038/s41596-020-0292-x

Ellenbroek, B., & Youn, J. (2016). Rodent models in neuroscience research: is it a rat race? *Disease models & mechanisms, 9*(10), 1079-1087. https://doi.org/10.1242/dmm.026120

Elmore, M. R. P., Najafi, A. R., Koike, M. A., Dagher, N. N., Spangenberg, E. E., Rice, R. A., Kitazawa, M., Matusow, B., Nguyen, H., West, B. L., & Green, K. N. (2014). Colony-stimulating factor 1 receptor signaling is necessary for microglia viability, unmasking a microglia progenitor cell in the adult brain. *Neuron, 82*(2), 380-397. https://doi.org/10.1016/j.neuron.2014.02.040

Engelhardt, B., Wolburg-Buchholz, K., & Wolburg, H. (2001, Jan 1). Involvement of the choroid plexus in central nervous system inflammation. *Microsc Res Tech, 52*(1), 112-129. https://doi.org/10.1002/1097-0029(20010101)52:1<112::Aid-jemt13>3.0.Co;2-5

Escartin, C., Guillemaud, O., & Carrillo-de Sauvage, M.-A. (2019, 2019/12/01). Questions and (some) answers on reactive astrocytes [https://doi.org/10.1002/glia.23687]. *Glia, 67*(12), 2221-2247. https://doi.org/https://doi.org/10.1002/glia.23687

Farina, C., Aloisi, F., & Meinl, E. (2007, 2007/03/01/). Astrocytes are active players in cerebral innate immunity. *Trends Immunol, 28*(3), 138-145. https://doi.org/https://doi.org/10.1016/j.it.2007.01.005

Fenn, A. M., Hall, J. C. E., Gensel, J. C., Popovich, P. G., & Godbout, J. P. (2014). IL-4 signaling drives a unique arginase+/IL-1β+ microglia phenotype and recruits macrophages to the inflammatory CNS: consequences of age-related deficits in IL-4Rα after traumatic spinal cord injury. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 34*(26), 8904-8917. https://doi.org/10.1523/JNEUROSCI.1146-14.2014

Field, C. S., Baixauli, F., Kyle, R. L., Puleston, D. J., Cameron, A. M., Sanin, D. E., Hippen, K. L., Loschi, M., Thangavelu, G., Corrado, M., Edwards-Hicks, J., Grzes, K. M., Pearce, E. J., Blazar, B. R., & Pearce, E. L. (2020, 2020/02/04/). Mitochondrial Integrity Regulated by Lipid Metabolism Is a Cell-Intrinsic Checkpoint for Treg Suppressive Function. *Cell Metabolism, 31*(2), 422-437.e425. https://doi.org/https://doi.org/10.1016/j.cmet.2019.11.021

Fleming, J. C., Norenberg, M. D., Ramsay, D. A., Dekaban, G. A., Marcillo, A. E., Saenz, A. D., Pasquale-Styles, M., Dietrich, W. D., & Weaver, L. C. (2006, Dec). The cellular inflammatory response in human spinal cords after injury. *Brain, 129*(Pt 12), 3249-3269. https://doi.org/10.1093/brain/awl296

Floriddia, E. M., Lourenço, T., Zhang, S., van Bruggen, D., Hilscher, M. M., Kukanja, P., Gonçalves Dos Santos, J. P., Altınkök, M., Yokota, C., Llorens-Bobadilla, E., Mulinyawe, S. B., Grãos, M., Sun, L. O., Frisén, J., Nilsson, M., & Castelo-Branco, G. (2020). Distinct oligodendrocyte populations have spatial preference and different responses to spinal cord injury. *Nat Commun, 11*(1), 5860-5860. https://doi.org/10.1038/s41467-020-19453-x

Fouad, K., Hurd, C., & Magnuson, D. S. K. (2013). Functional testing in animal models of spinal cord injury: not as straight forward as one would think. *Frontiers in integrative neuroscience, 7*, 85-85. https://doi.org/10.3389/fnint.2013.00085

Franzén, O., Gan, L.-M., & Björkegren, J. L. M. (2019). PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database, 2019*. https://doi.org/10.1093/database/baz046

Friedman, B. A., Srinivasan, K., Ayalon, G., Meilandt, W. J., Lin, H., Huntley, M. A., Cao, Y., Lee, S.-H., Haddick, P. C. G., Ngu, H., Modrusan, Z., Larson, J. L., Kaminker, J. S., van der Brug, M. P., & Hansen, D. V. (2018, 2018/01/16/). Diverse Brain Myeloid Expression Profiles Reveal Distinct Microglial Activation States and Aspects of Alzheimer's Disease Not Evident in Mouse Models. *Cell Rep, 22*(3), 832-847. https://doi.org/https://doi.org/10.1016/j.celrep.2017.12.066

Furuhashi, M., & Hotamisligil, G. S. (2008). Fatty acid-binding proteins: role in metabolic diseases and potential as drug targets. *Nature reviews. Drug discovery, 7*(6), 489-503. https://doi.org/10.1038/nrd2589

Gadani, Sachin P., Walsh, James T., Lukens, John R., & Kipnis, J. (2015). Dealing with Danger in the CNS: The Response of the Immune System to Injury. *Neuron, 87*(1), 47-62. https://doi.org/10.1016/j.neuron.2015.05.019

Gaudet, A. D., & Fonken, L. K. (2018, 2018/07/01). Glial Cells Shape Pathology and Repair After Spinal Cord Injury. *Neurotherapeutics, 15*(3), 554-577. https://doi.org/10.1007/s13311-018-0630-7

Gee, J. R., & Keller, J. N. (2005, 2005/06/01/). Astrocytes: regulation of brain homeostasis via apolipoprotein E. *The International Journal of Biochemistry & Cell Biology, 37*(6), 1145-1150. https://doi.org/https://doi.org/10.1016/j.biocel.2004.10.004

Gene Ontology Consortium: going forward. (2015, Jan). *Nucleic Acids Res, 43*(Database issue), D1049-1056. https://doi.org/10.1093/nar/gku1179

Gensel, J. C., & Zhang, B. (2015, 2015/09/04/). Macrophage activation and its role in repair and pathology after spinal cord injury. *Brain Research, 1619*, 1-11. https://doi.org/https://doi.org/10.1016/j.brainres.2014.12.045

Giesen, C., Wang, H. A. O., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P. J., Grolimund, D., Buhmann, J. M., Brandt, S., Varga, Z., Wild, P. J., Günther, D., & Bodenmiller, B. (2014, 2014/04/01). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods, 11*(4), 417-422. https://doi.org/10.1038/nmeth.2869

Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M. F., Conway, S. J., Ng, L. G., Stanley, E. R., Samokhvalov, I. M., & Merad, M. (2010). Fate Mapping

Analysis Reveals That Adult Microglia Derive from Primitive Macrophages. *Science, 330*(6005), 841. https://doi.org/10.1126/science.1194637

Goldmann, T., Wieghofer, P., Jordão, M. J. C., Prutek, F., Hagemeyer, N., Frenzel, K., Amann, L., Staszewski, O., Kierdorf, K., Krueger, M., Locatelli, G., Hochgerner, H., Zeiser, R., Epelman, S., Geissmann, F., Priller, J., Rossi, F. M. V., Bechmann, I., Kerschensteiner, M., Linnarsson, S., Jung, S., & Prinz, M. (2016). Origin, fate and dynamics of macrophages at central nervous system interfaces. *Nature Immunology, 17*(7), 797-805. https://doi.org/10.1038/ni.3423

Greenhalgh, A. D., & David, S. (2014, Apr 30). Differences in the phagocytic response of microglia and peripheral macrophages after spinal cord injury and its effects on cell death. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 34*(18), 6316-6322. https://doi.org/10.1523/jneurosci.4912-13.2014

Griffiths, J. A., Richard, A. C., Bach, K., Lun, A. T. L., & Marioni, J. C. (2018, 2018/07/10). Detection and removal of barcode swapping in single-cell RNA-seq data. *Nat Commun, 9*(1), 2667. https://doi.org/10.1038/s41467-018-05083-x

Griffiths, J. A., Scialdone, A., & Marioni, J. C. (2018, 2018/04/01). Using single-cell genomics to understand developmental processes and cell fate decisions [https://doi.org/10.15252/msb.20178046]. *Molecular Systems Biology, 14*(4), e8046. https://doi.org/https://doi.org/10.15252/msb.20178046

Guest, J. D., Hiester, E. D., & Bunge, R. P. (2005, 2005/04/01/). Demyelination and Schwann cell responses adjacent to injury epicenter cavities following chronic human spinal cord injury. *Experimental Neurology, 192*(2), 384-393. https://doi.org/https://doi.org/10.1016/j.expneurol.2004.11.033

Guilliams, M., Ginhoux, F., Jakubzick, C., Naik, S. H., Onai, N., Schraml, B. U., Segura, E., Tussiwand, R., & Yona, S. (2014, 2014/08/01). Dendritic cells, monocytes and macrophages: a unified nomenclature based on ontogeny. *Nature Reviews Immunology, 14*(8), 571-578. https://doi.org/10.1038/nri3712

Haage, V., Semtner, M., Vidal, R. O., Hernandez, D. P., Pong, W. W., Chen, Z., Hambardzumyan, D., Magrini, V., Ly, A., Walker, J., Mardis, E., Mertins, P., Sauer, S., Kettenmann, H., & Gutmann, D. H. (2019). Comprehensive gene expression meta-analysis identifies signature genes that distinguish microglia from peripheral monocytes/macrophages in health and glioma. *Acta Neuropathologica Communications, 7*(1), 20-20. https://doi.org/10.1186/s40478-019-0665-y

Hagemann-Jensen, M., Ziegenhain, C., Chen, P., Ramsköld, D., Hendriks, G.-J., Larsson, A. J. M., Faridani, O. R., & Sandberg, R. (2020, 2020/06/01). Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat Biotechnol, 38*(6), 708-714. https://doi.org/10.1038/s41587-020-0497-0

Haghverdi, L., Buettner, F., & Theis, F. J. (2015, Sep 15). Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics, 31*(18), 2989-2998. https://doi.org/10.1093/bioinformatics/btv325

Haghverdi, L., Lun, A. T. L., Morgan, M. D., & Marioni, J. C. (2018, Jun). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol, 36*(5), 421-427. https://doi.org/10.1038/nbt.4091

Haimon, Z., Volaski, A., Orthgiess, J., Boura-Halfon, S., Varol, D., Shemer, A., Yona, S., Zuckerman, B., David, E., Chappell-Maor, L., Bechmann, I., Gericke, M., Ulitsky, I., & Jung, S. (2018). Re-evaluating microglia expression profiles using RiboTag and cell isolation strategies. *Nature Immunology, 19*(6), 636-644. https://doi.org/10.1038/s41590-018-0110-6

Hamel, R., Peruzzotti-Jametti, L., Ridley, K., Testa, V., Yu, B., Rowitch, D., Marioni, J., & Pluchino, S. (2020). Time-resolved single-cell RNAseq profiling identifies a novel Fabp5-expressing subpopulation of inflammatory myeloid cells in chronic spinal cord injury. *bioRxiv*, 2020.2010.2021.346635. https://doi.org/10.1101/2020.10.21.346635

Hammond, T. R., Dufort, C., Dissing-Olesen, L., Giera, S., Young, A., Wysoker, A., Walker, A. J., Gergits, F., Segel, M., Nemesh, J., Marsh, S. E., Saunders, A., Macosko, E., Ginhoux, F., Chen, J., Franklin, R. J. M., Piao, X., McCarroll, S. A., & Stevens, B. (2019, Jan 15). Single-Cell RNA Sequencing of Microglia throughout the Mouse Lifespan and in the Injured Brain Reveals Complex Cell-State Changes. *Immunity, 50*(1), 253-271.e256. https://doi.org/10.1016/j.immuni.2018.11.004

Han, J., Fan, Y., Zhou, K., Blomgren, K., & Harris, R. A. (2021, 2021/03/17). Uncovering sex differences of rodent microglia. *J Neuroinflammation, 18*(1), 74. https://doi.org/10.1186/s12974-021-02124-z

Haque, A., Engel, J., Teichmann, S. A., & Lönnberg, T. (2017, 08/18). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Medicine, 9*, 75. https://doi.org/10.1186/s13073-017-0467-4

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K. J., Rozenblatt-Rosen, O., Dor, Y., Regev, A., & Yanai, I. (2016, Apr 28). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol, 17*, 77. https://doi.org/10.1186/s13059-016-0938-8

Hasselmann, J., Coburn, M. A., England, W., Figueroa Velez, D. X., Kiani Shabestari, S., Tu, C. H., McQuade, A., Kolahdouzan, M., Echeverria, K., Claes, C., Nakayama, T., Azevedo, R., Coufal, N. G., Han, C. Z., Cummings, B. J., Davtyan, H., Glass, C. K., Healy, L. M., Gandhi, S. P., Spitale, R. C., & Blurton-Jones, M. (2019). Development of a Chimeric Model to Study and Manipulate Human Microglia In Vivo [Article]. *Neuron, 103*(6), 1016-1033.e1010. https://doi.org/10.1016/j.neuron.2019.07.002

Hausmann, O. N. (2003, 2003/07/01). Post-traumatic inflammation following spinal cord injury. *Spinal Cord, 41*(7), 369-378. https://doi.org/10.1038/sj.sc.3101483

Hentschel, E., Will, C., Mustafi, N., Burkovski, A., Rehm, N., & Frunzke, J. (2013, 2013/03/01). Destabilized eYFP variants for dynamic gene expression studies in Corynebacterium glutamicum [https://doi.org/10.1111/j.1751-7915.2012.00360.x]. *Microbial Biotechnology, 6*(2), 196-201. https://doi.org/https://doi.org/10.1111/j.1751-7915.2012.00360.x

Hicks, S. C., Townes, F. W., Teng, M., & Irizarry, R. A. (2017, Nov 6). Missing data and technical variability in single-cell RNA-sequencing experiments. *Biostatistics*. https://doi.org/10.1093/biostatistics/kxx053

Hou, S., & Rabchevsky, A. G. (2014, Oct). Autonomic consequences of spinal cord injury. *Compr Physiol, 4*(4), 1419-1453. https://doi.org/10.1002/cphy.c130045

Hunter, C. A., & Jones, S. A. (2015, 2015/05/01). IL-6 as a keystone cytokine in health and disease. *Nature Immunology, 16*(5), 448-457. https://doi.org/10.1038/ni.3153

Inman, D., Guth, L., & Steward, O. (2002, Sep 23). Genetic influences on secondary degeneration and wound healing following spinal cord injury in various strains of mice. *The Journal of comparative neurology, 451*(3), 225-235. https://doi.org/10.1002/cne.10340

Iwata, A., Browne, K. D., Chen, X. H., Yuguchi, T., & Smith, D. H. (2005, Oct 1). Traumatic brain injury induces biphasic upregulation of ApoE and ApoJ protein in rats. *J Neurosci Res, 82*(1), 103-114. https://doi.org/10.1002/jnr.20607

Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., & Amit, I. (2014). Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types. *Science, 343*(6172), 776. https://doi.org/10.1126/science.1247651

Jakubzick, C. V., Randolph, G. J., & Henson, P. M. (2017, Jun). Monocyte differentiation and antigen-presenting functions. *Nat Rev Immunol, 17*(6), 349-362. https://doi.org/10.1038/nri.2017.28

James, S. L., Theadom, A., Ellenbogen, R. G., Bannick, M. S., Montjoy-Venning, W., Lucchesi, L. R., Abbasi, N., Abdulkader, R., Abraha, H. N., Adsuar, J. C., Afarideh, M., Agrawal, S., Ahmadi, A., Ahmed, M. B., Aichour, A. N., Aichour, I., Aichour, M. T. E., Akinyemi, R. O., Akseer, N., Alahdab, F., Alebel, A., Alghnam, S. A., Ali, B. A., Alsharif, U., Altirkawi, K., Andrei, C. L., Anjomshoa, M., Ansari, H., Ansha, M. G., Antonio, C. A. T., Appiah, S. C. Y., Ariani, F., Asefa, N. G., Asgedom, S. W., Atique, S., Awasthi, A., Ayala Quintanilla, B. P., Ayuk, T. B., Azzopardi, P. S., Badali, H., Badawi, A., Balalla, S., Banstola, A., Barker-Collo, S. L., Bärnighausen, T. W., Bedi, N., Behzadifar, M., Behzadifar, M., Bekele, B. B., Belachew, A. B., Belay, Y. A., Bennett, D. A., Bensenor, I. M., Berhane, A., Beuran, M., Bhalla, A., Bhaumik, S., Bhutta, Z. A., Biadgo, B., Biffino, M., Bijani, A., Bililign, N.,

Birungi, C., Boufous, S., Brazinova, A., Brown, A. W., Car, M., Cárdenas, R., Carrero, J. J., Carvalho, F., Castañeda-Orjuela, C. A., Catalá-López, F., Chaiah, Y., Champs, A. P., Chang, J.-C., Choi, J.-Y. J., Christopher, D. J., Cooper, C., Crowe, C. S., Dandona, L., Dandona, R., Daryani, A., Davitoiu, D. V., Degefa, M. G., Demoz, G. T., Deribe, K., Djalalinia, S., Do, H. P., Doku, D. T., Drake, T. M., Dubey, M., Dubljanin, E., El-Khatib, Z., Ofori-Asenso, R., Eskandarieh, S., Esteghamati, A., Esteghamati, S., Faro, A., Farzadfar, F., Farzaei, M. H., Fereshtehnejad, S.-M., Fernandes, E., Feyissa, G. T., Filip, I., Fischer, F., Fukumoto, T., Ganji, M., Gankpe, F. G., Gebre, A. K., Gebrehiwot, T. T., Gezae, K. E., Gopalkrishna, G., Goulart, A. C., Haagsma, J. A., Haj-Mirzaian, A., Haj-Mirzaian, A., Hamadeh, R. R., Hamidi, S., Haro, J. M., Hassankhani, H., Hassen, H. Y., Havmoeller, R., Hawley, C., Hay, S. I., Hegazy, M. I., Hendrie, D., Henok, A., Hibstu, D. T., Hoffman, H. J., Hole, M. K., Homaie Rad, E., Hosseini, S. M., Hostiuc, S., Hu, G., Hussen, M. A., Ilesanmi, O. S., Irvani, S. S. N., Jakovljevic, M., Jayaraman, S., Jha, R. P., Jonas, J. B., Jones, K. M., Jorjoran Shushtari, Z., Jozwiak, J. J., Jürisson, M., Kabir, A., Kahsay, A., Kahssay, M., Kalani, R., Karch, A., Kasaeian, A., Kassa, G. M., Kassa, T. D., Kassa, Z. Y., Kengne, A. P., Khader, Y. S., Khafaie, M. A., Khalid, N., Khalil, I., Khan, E. A., Khan, M. S., Khang, Y.-H., Khazaie, H., Khoja, A. T., Khubchandani, J., Kiadaliri, A. A., Kim, D., Kim, Y.-E., Kisa, A., Koyanagi, A., Krohn, K. J., Kuate Defo, B., Kucuk Bicer, B., Kumar, G. A., Kumar, M., Lalloo, R., Lami, F. H., Lansingh, V. C., Laryea, D. O., Latifi, A., Leshargie, C. T., Levi, M., Li, S., Liben, M. L., Lotufo, P. A., Lunevicius, R., Mahotra, N. B., Majdan, M., Majeed, A., Malekzadeh, R., Manda, A.-L., Mansournia, M. A., Massenburg, B. B., Mate, K. K. V., Mehndiratta, M. M., Mehta, V., Meles, H., Melese, A., Memiah, P. T. N., Mendoza, W., Mengistu, G., Meretoja, A., Meretoja, T. J., Mestrovic, T., Miazgowski, T., Miller, T. R., Mini, G. K., Mirica, A., Mirrakhimov, E. M., Moazen, B., Mohammadi, M., Molokhia, M., Monasta, L., Mondello, S., Moosazadeh, M., Moradi, G., Moradi, M., Moradi-Lakeh, M., Moradinazar, M., Morrison, S. D., Moschos, M. M., Mousavi, S. M., Murthy, S., Musa, K. I., Mustafa, G., Naghavi, M., Naik, G., Najafi, F., Nangia, V., Nascimento, B. R., Negoi, I., Nguyen, T. H., Nichols, E., Ningrum, D. N. A., Nirayo, Y. L., Nyasulu, P. S., Ogbo, F. A., Oh, I.-H., Okoro, A., Olagunju, A. T., Olagunju, T. O., Olivares, P. R., Otstavnov, S. S., Owolabi, M. O., P A, M., Pakhale, S., Pandey, A. R., Pesudovs, K., Pinilla-Monsalve, G. D., Polinder, S., Poustchi, H., Prakash, S., Qorbani, M., Radfar, A., Rafay, A., Rafiei, A., Rahimi-Movaghar, A., Rahimi-Movaghar, V., Rahman, M., Rahman, M. A., Rai, R. K., Rajati, F., Ram, U., Rawaf, D. L., Rawaf, S., Reiner, R. C., Reis, C., Renzaho, A. M. N., Resnikoff, S., Rezaei, S., Rezaeian, S., Roever, L., Ronfani, L., Roshandel, G., Roy, N., Ruhago, G. M., Saddik, B., Safari, H., Safiri, S., Sahraian, M. A., Salamati, P., Saldanha, R. d. F., Samy, A. M., Sanabria, J., Santos, J. V., Santric Milicevic, M. M. M., Sartorius, B., Satpathy, M., Savuon, K., Schneider, I. J. C., Schwebel, D. C., Sepanlou, S. G., Shabaninejad, H., Shaikh, M. A. A., Shams-Beyranvand, M., Sharif, M., Sharif-Alhoseini, M., Shariful Islam, S. M., She, J., Sheikh, A., Shen, J., Sheth, K. N., Shibuya, K., Shiferaw, M. S., Shigematsu, M., Shiri, R., Shiue, I., Shoman, H., Siabani, S., Siddiqi, T. J., Silva, J. P., Silveira, D. G. A., Sinha, D. N., Smith, M., Soares Filho, A. M., Sobhani, S., Soofi, M., Soriano, J. B., Soyiri, I. N., Stein, D. J., Stokes, M. A., Sufiyan, M. a. B., Sunguya, B. F., Sunshine, J. E., Sykes, B. L., Szoeke, C. E. I., Tabarés-Seisdedos, R., Te Ao, B. J., Tehrani-Banihashemi, A., Tekle, M. G., Temsah, M.-H., Temsah, O., Topor-Madry, R., Tortajada-Girbés, M., Tran, B. X., Tran, K. B., Tudor Car, L., Ukwaja, K. N., Ullah, I., Usman, M. S., Uthman, O. A., Valdez, P. R., Vasankari, T. J., Venketasubramanian, N., Violante, F. S.,

Wagnew, F. W. S., Waheed, Y., Wang, Y.-P., Weldegwergs, K. G., Werdecker, A., Wijeratne, T., Winkler, A. S., Wyper, G. M. A., Yano, Y., Yaseri, M., Yasin, Y. J., Ye, P., Yimer, E. M., Yip, P., Yisma, E., Yonemoto, N., Yoon, S.-J., Yost, M. G., Younis, M. Z., Yousefifard, M., Yu, C., Zaidi, Z., Zaman, S. B., Zamani, M., Zenebe, Z. M., Zodpey, S., Feigin, V. L., Vos, T., & Murray, C. J. L. (2019). Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology, 18*(1), 56-87. https://doi.org/10.1016/S1474-4422(18)30415-0

Jiang, S. X., Slinn, J., Aylsworth, A., & Hou, S. T. (2012, Aug). Vimentin participates in microglia activation and neurotoxicity in cerebral ischemia. *J Neurochem, 122*(4), 764-774. https://doi.org/10.1111/j.1471-4159.2012.07823.x

Jones, T. B. (2014, 2014/08/01/). Lymphocytes and autoimmunity after spinal cord injury. *Experimental Neurology, 258*, 78-90. https://doi.org/https://doi.org/10.1016/j.expneurol.2014.03.003

Jordão, M. J. C., Sankowski, R., Brendecke, S. M., Sagar, Locatelli, G., Tai, Y.-H., Tay, T. L., Schramm, E., Armbruster, S., Hagemeyer, N., Groß, O., Mai, D., Çiçek, Ö., Falk, T., Kerschensteiner, M., Grün, D., & Prinz, M. (2019). Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. *Science, 363*(6425), eaat7554. https://doi.org/10.1126/science.aat7554

Kaczocha, M., Rebecchi, M. J., Ralph, B. P., Teng, Y.-H. G., Berger, W. T., Galbavy, W., Elmes, M. W., Glaser, S. T., Wang, L., Rizzo, R. C., Deutsch, D. G., & Ojima, I. (2014). Inhibition of fatty acid binding proteins elevates brain anandamide levels and produces analgesia. *PloS one, 9*(4), e94200-e94200. https://doi.org/10.1371/journal.pone.0094200

Kakebeen, A. D., Chitsazan, A. D., Williams, M. C., Saunders, L. M., & Wills, A. E. (2020, 2020/04/27). Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors. *Elife, 9*, e52648. https://doi.org/10.7554/eLife.52648

Kannan-Thulasiraman, P., Seachrist, D. D., Mahabeleshwar, G. H., Jain, M. K., & Noy, N. (2010, Jun 18). Fatty acid-binding protein 5 and PPARbeta/delta are critical mediators of epidermal growth factor receptor-induced carcinoma cell growth. *The Journal of biological chemistry, 285*(25), 19106-19115. https://doi.org/10.1074/jbc.M109.099770

Keren-Shaul, H., Spinrad, A., Weiner, A., Matcovitch-Natan, O., Dvir-Szternfeld, R., Ulland, T. K., David, E., Baruch, K., Lara-Astaiso, D., Toth, B., Itzkovitz, S., Colonna, M., Schwartz, M., & Amit, I. (2017, 2017/06/15/). A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell, 169*(7), 1276-1290.e1217. https://doi.org/https://doi.org/10.1016/j.cell.2017.05.018

Kezic, J., & McMenamin, P. G. (2008, 2008/09/01). Differential turnover rates of monocyte-derived cells in varied ocular tissue microenvironments. *Journal of Leukocyte Biology, 84*(3), 721-729. https://doi.org/10.1189/jlb.0308166

Kierdorf, K., Erny, D., Goldmann, T., Sander, V., Schulz, C., Perdiguero, E. G., Wieghofer, P., Heinrich, A., Riemke, P., Hölscher, C., Müller, D. N., Luckow, B., Brocker, T., Debowski, K., Fritz, G., Opdenakker, G., Diefenbach, A., Biber, K., Heikenwalder, M., Geissmann, F., Rosenbauer, F., & Prinz, M. (2013, 2013/03/01). Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nature Neuroscience, 16*(3), 273-280. https://doi.org/10.1038/nn.3318

Kierdorf, K., Masuda, T., Jordão, M. J. C., & Prinz, M. (2019, 2019/09/01). Macrophages at CNS interfaces: ontogeny and function in health and disease. *Nature Reviews Neuroscience, 20*(9), 547-562. https://doi.org/10.1038/s41583-019-0201-x

Kierdorf, K., & Prinz, M. (2017, 09/01/). Microglia in steady state. *The Journal of clinical investigation, 127*(9), 3201-3209. https://doi.org/10.1172/JCI90602

Kigerl, K. A., Gensel, J. C., Ankeny, D. P., Alexander, J. K., Donnelly, D. J., & Popovich, P. G. (2009). Identification of two distinct macrophage subsets with divergent effects causing either neurotoxicity or regeneration in the injured mouse spinal cord. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 29*(43), 13435-13444. https://doi.org/10.1523/JNEUROSCI.3257-09.2009

Kim, C. C., Nakamura, M. C., & Hsieh, C. L. (2016, May 24). Brain trauma elicits non-canonical macrophage activation states. *J Neuroinflammation, 13*(1), 117. https://doi.org/10.1186/s12974-016-0581-z

Kim, E., Woo, M. S., Qin, L., Ma, T., Beltran, C. D., Bao, Y., Bailey, J. A., Corbett, D., Ratan, R. R., Lahiri, D. K., & Cho, S. (2015, Nov 11). Daidzein Augments Cholesterol Homeostasis via ApoE to Promote Functional Recovery in Chronic Stroke. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 35*(45), 15113-15126. https://doi.org/10.1523/jneurosci.2890-15.2015

Kivisäkk, P., Imitola, J., Rasmussen, S., Elyaman, W., Zhu, B., Ransohoff, R. M., & Khoury, S. J. (2009, 2009/04/01). Localizing central nervous system immune surveillance: Meningeal antigen-presenting cells activate T cells during experimental autoimmune encephalomyelitis [https://doi.org/10.1002/ana.21379]. *Annals of Neurology, 65*(4), 457-469. https://doi.org/https://doi.org/10.1002/ana.21379

Klein, Allon M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, David A., & Kirschner, Marc W. (2015, 2015/05/21/). Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell, 161*(5), 1187-1201. https://doi.org/https://doi.org/10.1016/j.cell.2015.04.044

Kong, F.-Q., Zhao, S.-J., Sun, P., Liu, H., Jie, J., Xu, T., Xu, A.-D., Yang, Y.-Q., Zhu, Y., Chen, J., Zhou, Z., Qian, D.-F., Gu, C.-J., Chen, Q., Yin, G.-Y., Zhang, H.-W., & Fan, J. (2020,

2020/02/17). Macrophage MSR1 promotes the formation of foamy macrophage and neuronal apoptosis after spinal cord injury. *J Neuroinflammation, 17*(1), 62. https://doi.org/10.1186/s12974-020-01735-2

Kong, X., & Gao, J. (2017, May). Macrophage polarization: a key event in the secondary phase of acute spinal cord injury. *J Cell Mol Med, 21*(5), 941-954. https://doi.org/10.1111/jcmm.13034

Konishi, H., Kobayashi, M., Kunisawa, T., Imai, K., Sayo, A., Malissen, B., Crocker, P. R., Sato, K., & Kiyama, H. (2017, Dec). Siglec-H is a microglia-specific marker that discriminates microglia from CNS-associated macrophages and CNS-infiltrating monocytes. *Glia, 65*(12), 1927-1943. https://doi.org/10.1002/glia.23204

Kopper, T. J., & Gensel, J. C. (2018). Myelin as an inflammatory mediator: Myelin interactions with complement, macrophages, and microglia in spinal cord injury. *Journal of Neuroscience Research, 96*(6), 969-977. https://doi.org/10.1002/jnr.24114

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-r., & Raychaudhuri, S. (2019, 2019/12/01). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods, 16*(12), 1289-1296. https://doi.org/10.1038/s41592-019-0619-0

Krasemann, S., Madore, C., Cialic, R., Baufeld, C., Calcagno, N., El Fatimy, R., Beckers, L., O'Loughlin, E., Xu, Y., Fanek, Z., Greco, D. J., Smith, S. T., Tweet, G., Humulock, Z., Zrzavy, T., Conde-Sanroman, P., Gacias, M., Weng, Z., Chen, H., Tjon, E., Mazaheri, F., Hartmann, K., Madi, A., Ulrich, J. D., Glatzel, M., Worthmann, A., Heeren, J., Budnik, B., Lemere, C., Ikezu, T., Heppner, F. L., Litvak, V., Holtzman, D. M., Lassmann, H., Weiner, H. L., Ochando, J., Haass, C., & Butovsky, O. (2017). The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity, 47*(3), 566-581.e569. https://doi.org/10.1016/j.immuni.2017.08.008

Kubick, N., Henckell Flournoy, P. C., Klimovich, P., Manda, G., & Mickael, M.-E. (2020, 2020/12/01). What has single-cell RNA sequencing revealed about microglial neuroimmunology? [https://doi.org/10.1002/iid3.362]. *Immunity, Inflammation and Disease, 8*(4), 825-839. https://doi.org/https://doi.org/10.1002/iid3.362

Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D. J., Hicks, S. C., Robinson, M. D., Vallejos, C. A., Campbell, K. R., Beerenwinkel, N., Mahfouz, A., Pinello, L., Skums, P., Stamatakis, A., Attolini, C. S.-O., Aparicio, S., Baaijens, J., Balvert, M., Barbanson, B. d., Cappuccio, A., Corleone, G., Dutilh, B. E., Florescu, M., Guryev, V., Holmer, R., Jahn, K., Lobo, T. J., Keizer, E. M., Khatri, I., Kielbasa, S. M., Korbel, J. O., Kozlov, A. M., Kuo, T.-H., Lelieveldt, B. P. F., Mandoiu, I. I., Marioni, J. C., Marschall, T., Mölder, F., Niknejad, A., Raczkowski, L., Reinders, M., Ridder, J. d., Saliba, A.-E., Somarakis, A., Stegle, O., Theis, F. J., Yang, H., Zelikovsky, A., McHardy, A. C., Raphael, B. J., Shah, S. P., & Schönhuth, A. (2020, 2020/02/07). Eleven grand challenges in single-cell data science. *Genome Biology, 21*(1), 31. https://doi.org/10.1186/s13059-020-1926-6

Lee-Liu, D., Edwards-Faret, G., Tapia, V. S., & Larraín, J. (2013, Aug). Spinal cord regeneration: lessons for mammals from non-mammalian vertebrates. *Genesis, 51*(8), 529-544. https://doi.org/10.1002/dvg.22406

Levi, L., Lobo, G., Doud, M. K., von Lintig, J., Seachrist, D., Tochtrop, G. P., & Noy, N. (2013, Aug 1). Genetic ablation of the fatty acid-binding protein FABP5 suppresses HER2-induced mammary tumorigenesis. *Cancer Res, 73*(15), 4770-4780. https://doi.org/10.1158/0008-5472.Can-13-0384

Li, J. J., Liu, H., Zhu, Y., Yan, L., Liu, R., Wang, G., Wang, B., & Zhao, B. (2020). Animal Models for Treating Spinal Cord Injury Using Biomaterials-Based Tissue Engineering Strategies. *Tissue Engineering Part B: Reviews*. https://doi.org/10.1089/ten.teb.2020.0267

Li, Q., & Barres, B. A. (2018, 2018/04/01). Microglia and macrophages in brain homeostasis and disease. *Nature Reviews Immunology, 18*(4), 225-242. https://doi.org/10.1038/nri.2017.125

Li, Q., Cheng, Z., Zhou, L., Darmanis, S., Neff, N. F., Okamoto, J., Gulati, G., Bennett, M. L., Sun, L. O., Clarke, L. E., Marschallinger, J., Yu, G., Quake, S. R., Wyss-Coray, T., & Barres, B. A. (2019, Jan 16). Developmental Heterogeneity of Microglia and Brain Myeloid Cells Revealed by Deep Single-Cell RNA Sequencing. *Neuron, 101*(2), 207-223.e210. https://doi.org/10.1016/j.neuron.2018.12.006

Li, Y., He, X., Kawaguchi, R., Zhang, Y., Wang, Q., Monavarfeshani, A., Yang, Z., Chen, B., Shi, Z., Meng, H., Zhou, S., Zhu, J., Jacobi, A., Swarup, V., Popovich, P. G., Geschwind, D. H., & He, Z. (2020, 2020/10/07). Microglia-organized scar-free spinal cord repair in neonatal mice. *Nature*. https://doi.org/10.1038/s41586-020-2795-6

Liddelow, S. A., & Barres, B. A. (2017, Jun 20). Reactive Astrocytes: Production, Function, and Therapeutic Potential. *Immunity, 46*(6), 957-967. https://doi.org/10.1016/j.immuni.2017.06.006

Lilley, E., Andrews, M. R., Bradbury, E. J., Elliott, H., Hawkins, P., Ichiyama, R. M., Keeley, J., Michael-Titus, A. T., Moon, L. D. F., Pluchino, S., Riddell, J., Ryder, K., & Yip, P. K. (2020, 2020/06/01/). Refining rodent models of spinal cord injury. *Experimental Neurology, 328*, 113273. https://doi.org/https://doi.org/10.1016/j.expneurol.2020.113273

Liu, J., Qian, C., & Cao, X. (2016, Jul 19). Post-Translational Modification Control of Innate Immunity. *Immunity, 45*(1), 15-30. https://doi.org/10.1016/j.immuni.2016.06.020

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell, 165*(3), 535-550. https://doi.org/10.1016/j.cell.2016.03.014

Liu, Y., Yang, M., Deng, Y., Su, G., Enninful, A., Guo, C. C., Tebaldi, T., Zhang, D., Kim, D., Bai, Z., Norris, E., Pan, A., Li, J., Xiao, Y., Halene, S., & Fan, R. (2020). High-Spatial-Resolution Multi-Omics Sequencing via Deterministic Barcoding in Tissue. *Cell, 183*(6), 1665-1681.e1618. https://doi.org/10.1016/j.cell.2020.10.026

Loving, B. A., & Bruce, K. D. (2020, 2020-April-28). Lipid and Lipoprotein Metabolism in Microglia [Review]. *Frontiers in Physiology, 11*(393). https://doi.org/10.3389/fphys.2020.00393

Luecken, M. D., & Theis, F. J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology, 15*(6), e8746. https://doi.org/10.15252/msb.20188746

Lun, A. M., DJ; Marioni, JC. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000 Research, 5*(2122). https://doi.org/10.12688/f1000research.9501.2

Lun, A. T. L., Riesenfeld, S., Andrews, T., Dao, T. P., Gomes, T., Marioni, J. C., & participants in the 1st Human Cell Atlas, J. (2019, 2019/03/22). EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biology, 20*(1), 63. https://doi.org/10.1186/s13059-019-1662-y

Madisen, L., Zwingman, T. A., Sunkin, S. M., Oh, S. W., Zariwala, H. A., Gu, H., Ng, L. L., Palmiter, R. D., Hawrylycz, M. J., Jones, A. R., Lein, E. S., & Zeng, H. (2010, Jan). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci, 13*(1), 133-140. https://doi.org/10.1038/nn.2467

Mamanova, L., Miao, Z., Jinat, A., Ellis, P., Shirley, L., & Teichmann, S. A. (2021, 2021/06/01). High-throughput full-length single-cell RNA-seq automation. *Nature Protocols, 16*(6), 2886-2915. https://doi.org/10.1038/s41596-021-00523-3

Marinelli, S., Vacca, V., De Angelis, F., Pieroni, L., Orsini, T., Parisi, C., Soligo, M., Protto, V., Manni, L., Guerrieri, R., & Pavone, F. (2019, 2019/06/20). Innovative mouse model mimicking human-like features of spinal cord injury: efficacy of Docosahexaenoic acid on acute and chronic phases. *Scientific Reports, 9*(1), 8883. https://doi.org/10.1038/s41598-019-45037-x

Marinov, G. K., Williams, B. A., McCue, K., Schroth, G. P., Gertz, J., Myers, R. M., & Wold, B. J. (2014). From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome research, 24*(3), 496-510. https://doi.org/10.1101/gr.161034.113

Marschallinger, J., Iram, T., Zardeneta, M., Lee, S. E., Lehallier, B., Haney, M. S., Pluvinage, J. V., Mathur, V., Hahn, O., Morgens, D. W., Kim, J., Tevini, J., Felder, T. K., Wolinski, H., Bertozzi, C. R., Bassik, M. C., Aigner, L., & Wyss-Coray, T. (2020, 2020/01/20). Lipid-droplet-accumulating microglia represent a dysfunctional and proinflammatory state

in the aging brain. *Nature Neuroscience*. https://doi.org/10.1038/s41593-019-0566-1

Marsh, S. E., Kamath, T., Walker, A. J., Dissing-Olesen, L., Hammond, T. R., Young, A. M. H., Abdulraouf, A., Nadaf, N., Dufort, C., Murphy, S., Kozareva, V., Vanderburg, C., Hong, S., Bulstrode, H., Hutchinson, P. J., Gaffney, D. J., Franklin, R. J. M., Macosko, E. Z., & Stevens, B. (2020). Single Cell Sequencing Reveals Glial Specific Responses to Tissue Processing & Enzymatic Dissociation in Mice and Humans. *bioRxiv*, 2020.2012.2003.408542. https://doi.org/10.1101/2020.12.03.408542

Masuda, T., Sankowski, R., Staszewski, O., & Prinz, M. (2020, 2020/02/04/). Microglia Heterogeneity in the Single-Cell Era. *Cell Rep, 30*(5), 1271-1281. https://doi.org/https://doi.org/10.1016/j.celrep.2020.01.010

Matcovitch-Natan, O., Winter, D. R., Giladi, A., Vargas Aguilar, S., Spinrad, A., Sarrazin, S., Ben-Yehuda, H., David, E., Zelada González, F., Perrin, P., Keren-Shaul, H., Gury, M., Lara-Astaiso, D., Thaiss, C. A., Cohen, M., Bahar Halpern, K., Baruch, K., Deczkowska, A., Lorenzo-Vivas, E., Itzkovitz, S., Elinav, E., Sieweke, M. H., Schwartz, M., & Amit, I. (2016). Microglia development follows a stepwise program to regulate brain homeostasis. *Science, 353*(6301), aad8670. https://doi.org/10.1126/science.aad8670

May, P., Reddy, Y. K., & Herz, J. (2002, May 24). Proteolytic processing of low density lipoprotein receptor-related protein mediates regulated release of its intracellular domain. *The Journal of biological chemistry, 277*(21), 18736-18743. https://doi.org/10.1074/jbc.M201979200

McCarthy, D. J., Campbell, K. R., Lun, A. T. L., & Wills, Q. F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics, 33*(8), 1179-1186. https://doi.org/10.1093/bioinformatics/btw777

McGinnis, C. S., Patterson, D. M., Winkler, J., Conrad, D. N., Hein, M. Y., Srivastava, V., Hu, J. L., Murrow, L. M., Weissman, J. S., Werb, Z., Chow, E. D., & Gartner, Z. J. (2019, Jul). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods, 16*(7), 619-626. https://doi.org/10.1038/s41592-019-0433-8

Meiser, J., Krämer, L., Sapcariu, S. C., Battello, N., Ghelfi, J., D'Herouel, A. F., Skupin, A., & Hiller, K. (2016). Pro-inflammatory Macrophages Sustain Pyruvate Oxidation through Pyruvate Dehydrogenase for the Synthesis of Itaconate and to Enable Cytokine Expression *. *Journal of Biological Chemistry, 291*(8), 3932-3946. https://doi.org/10.1074/jbc.M115.676817

Mestas, J., & Hughes, C. C. W. (2004). Of Mice and Not Men: Differences between Mouse and Human Immunology. *The Journal of Immunology, 172*(5), 2731. https://doi.org/10.4049/jimmunol.172.5.2731

Mi, H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013, Aug). Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc, 8*(8), 1551-1566. https://doi.org/10.1038/nprot.2013.092

Mildner, A., Schlevogt, B., Kierdorf, K., Böttcher, C., Erny, D., Kummer, M. P., Quinn, M., Brück, W., Bechmann, I., Heneka, M. T., Priller, J., & Prinz, M. (2011). Distinct and Non-Redundant Roles of Microglia and Myeloid Subsets in Mouse Models of Alzheimer's Disease. *The Journal of Neuroscience, 31*(31), 11159. https://doi.org/10.1523/JNEUROSCI.6209-10.2011

Mildner, A., Schmidt, H., Nitsche, M., Merkler, D., Hanisch, U. K., Mack, M., Heikenwalder, M., Brück, W., Priller, J., & Prinz, M. (2007, Dec). Microglia in the adult brain arise from Ly-6ChiCCR2+ monocytes only under defined host conditions. *Nat Neurosci, 10*(12), 1544-1553. https://doi.org/10.1038/nn2015

Milich, L. M., Choi, J. S., Ryan, C., Cerqueira, S. R., Benavides, S., Yahn, S. L., Tsoulfas, P., & Lee, J. K. (2021). Single-cell analysis of the cellular heterogeneity and interactions in the injured mouse spinal cord. *Journal of Experimental Medicine, 218*(8). https://doi.org/10.1084/jem.20210040

Milich, L. M., Ryan, C. B., & Lee, J. K. (2019, 2019/05/01). The origin, fate, and contribution of macrophages to spinal cord injury pathology. *Acta Neuropathologica, 137*(5), 785-797. https://doi.org/10.1007/s00401-019-01992-3

Moore, S. M., Holt, V. V., Malpass, L. R., Hines, I. N., & Wheeler, M. D. (2015, Oct). Fatty acid-binding protein 5 limits the anti-inflammatory response in murine macrophages. *Mol Immunol, 67*(2 Pt B), 265-275. https://doi.org/10.1016/j.molimm.2015.06.001

Morris, S. A. (2019). The evolving concept of cell identity in the single cell era. *Development, 146*(12). https://doi.org/10.1242/dev.169748

Müller, C., Bauer, N. M., Schäfer, I., & White, R. (2013). Making myelin basic protein -from mRNA transport to localized translation. *Frontiers in cellular neuroscience, 7*, 169-169. https://doi.org/10.3389/fncel.2013.00169

Nahrendorf, M., & Swirski, F. K. (2016, 2016/07/22). Abandoning M1/M2 for a Network Model of Macrophage Function. *Circulation Research, 119*(3), 414-417. https://doi.org/10.1161/CIRCRESAHA.116.309194

Nayak, R., & Hasija, Y. (2021, 2021/03/01/). A hitchhiker's guide to single-cell transcriptomics and data analysis pipelines. *Genomics, 113*(2), 606-619. https://doi.org/https://doi.org/10.1016/j.ygeno.2021.01.007

Neirinckx, V., Coste, C., Franzen, R., Gothot, A., Rogister, B., & Wislet, S. (2014). Neutrophil contribution to spinal cord injury and repair. *J Neuroinflammation, 11*, 150-150. https://doi.org/10.1186/s12974-014-0150-2

Nishimura, S., Sasaki, T., Shimizu, A., Yoshida, K., Iwai, H., Koya, I., Kobayashi, Y., Itakura, G., Shibata, S., Ebise, H., Horiuchi, K., Kudoh, J., Toyama, Y., Anderson, A. J., Okano, H., & Nakamura, M. (2014, 2014/11/01/). Global gene expression analysis following spinal cord injury in non-human primates. *Experimental Neurology, 261*, 171-179. https://doi.org/https://doi.org/10.1016/j.expneurol.2014.05.021

Noristani, H. N., Gerber, Y. N., Sabourin, J.-C., Le Corre, M., Lonjon, N., Mestre-Frances, N., Hirbec, H. E., & Perrin, F. E. (2017, 2017-April-03). RNA-Seq Analysis of Microglia Reveals Time-Dependent Activation of Specific Genetic Programs following Spinal Cord Injury [Original Research]. *Frontiers in Molecular Neuroscience, 10*(90). https://doi.org/10.3389/fnmol.2017.00090

Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLOS Biology, 18*(3), e3000691. https://doi.org/10.1371/journal.pbio.3000691

Ochocka, N., Segit, P., Walentynowicz, K. A., Wojnicki, K., Cyranowski, S., Swatler, J., Mieczkowski, J., & Kaminska, B. (2021, 2021/02/19). Single-cell RNA sequencing reveals functional heterogeneity of glioma-associated brain macrophages. *Nat Commun, 12*(1), 1151. https://doi.org/10.1038/s41467-021-21407-w

Okada, S., Hara, M., Kobayakawa, K., Matsumoto, Y., & Nakashima, Y. (2018, Jan). Astrocyte reactivity and astrogliosis after spinal cord injury. *Neurosci Res, 126*, 39-43. https://doi.org/10.1016/j.neures.2017.10.004

Okada, S., Nakamura, M., Katoh, H., Miyao, T., Shimazaki, T., Ishii, K., Yamane, J., Yoshimura, A., Iwamoto, Y., Toyama, Y., & Okano, H. (2006, Jul). Conditional ablation of Stat3 or Socs3 discloses a dual role for reactive astrocytes after spinal cord injury. *Nat Med, 12*(7), 829-834. https://doi.org/10.1038/nm1425

Orr, M. B., & Gensel, J. C. (2018, 2018/07/01). Spinal Cord Injury Scarring and Inflammation: Therapies Targeting Glial and Inflammatory Responses. *Neurotherapeutics, 15*(3), 541-553. https://doi.org/10.1007/s13311-018-0631-6

Osorio, D., & Cai, J. J. (2021). Systematic determination of the mitochondrial proportion in human and mice tissues for single-cell RNA-sequencing data quality control. *Bioinformatics, 37*(7), 963-967. https://doi.org/10.1093/bioinformatics/btaa751

Osterburg, A. R., Hexley, P., Supp, D. M., Robinson, C. T., Noel, G., Ogle, C., Boyce, S. T., Aronow, B. J., & Babcock, G. F. (2013). Concerns over interspecies transcriptional comparisons in mice and humans after trauma. *Proceedings of the National Academy of Sciences of the United States of America, 110*(36), E3370-E3370. https://doi.org/10.1073/pnas.1306033110

Papalexi, E., & Satija, R. (2018, 2018/01/01). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nature Reviews Immunology, 18*(1), 35-45. https://doi.org/10.1038/nri.2017.76

Parkhurst, C. N., Yang, G., Ninan, I., Savas, J. N., Yates, J. R., 3rd, Lafaille, J. J., Hempstead, B. L., Littman, D. R., & Gan, W. B. (2013, Dec 19). Microglia promote learning-dependent synapse formation through brain-derived neurotrophic factor. *Cell, 155*(7), 1596-1609. https://doi.org/10.1016/j.cell.2013.11.030

Penkowa, M., Giralt, M., Lago, N., Camats, J., Carrasco, J., Hernández, J., Molinero, A., Campbell, I. L., & Hidalgo, J. (2003, Jun). Astrocyte-targeted expression of IL-6 protects the CNS against a focal brain injury. *Exp Neurol, 181*(2), 130-148. https://doi.org/10.1016/s0014-4886(02)00051-1

Peruzzotti-Jametti, L., Willis, C. M., Hamel, R., Krzak, G., & Pluchino, S. (2021, 2021-June-23). Metabolic Control of Smoldering Neuroinflammation [Review]. *Frontiers in Immunology, 12*(2340). https://doi.org/10.3389/fimmu.2021.705920

Peterson, V. M., Zhang, K. X., Kumar, N., Wong, J., Li, L., Wilson, D. C., Moore, R., McClanahan, T. K., Sadekova, S., & Klappenbach, J. A. (2017, Oct). Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol, 35*(10), 936-939. https://doi.org/10.1038/nbt.3973

Phillips, R. J., Lutz, M., & Premack, B. (2005, 2005/10/31). Differential signaling mechanisms regulate expression of CC chemokine receptor-2 during monocyte maturation. *Journal of Inflammation, 2*(1), 14. https://doi.org/10.1186/1476-9255-2-14

Picelli, S. (2017, 07/21
03/30/received
06/08/revised
06/09/accepted). Single-cell RNA-sequencing: The future of genome biology is now. *RNA Biology, 14*(5), 637-650. https://doi.org/10.1080/15476286.2016.1201618

Plemel, J. R., Stratton, J. A., Michaels, N. J., Rawji, K. S., Zhang, E., Sinha, S., Baaklini, C. S., Dong, Y., Ho, M., Thorburn, K., Friedman, T. N., Jawad, S., Silva, C., Caprariello, A. V., Hoghooghi, V., Yue, J., Jaffer, A., Lee, K., Kerr, B. J., Midha, R., Stys, P. K., Biernaskie, J., & Yong, V. W. (2020). Microglia response following acute demyelination is heterogeneous and limits infiltrating macrophage dispersion. *Science Advances, 6*(3), eaay6324. https://doi.org/10.1126/sciadv.aay6324

Popovich, P. G., & Jones, T. B. (2003, Jan). Manipulating neuroinflammatory reactions in the injured spinal cord: back to basics. *Trends Pharmacol Sci, 24*(1), 13-17. https://doi.org/10.1016/s0165-6147(02)00006-8

Prinz, M., Erny, D., & Hagemeyer, N. (2017, 03/22/online). Ontogeny and homeostasis of CNS myeloid cells [Review Article]. *Nature Immunology, 18*, 385. https://doi.org/10.1038/ni.3703

Qi, M., Li, W., Tsang, I. W., & Yijun, S. (2017). Principal Graph and Structure Learning Based on Reversed Graph Embedding. *IEEE transactions on pattern analysis and machine intelligence, 39*(11), 2227-2241. https://doi.org/10.1109/TPAMI.2016.2635657

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., & Trapnell, C. (2017, Oct). Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods, 14*(10), 979-982. https://doi.org/10.1038/nmeth.4402

Rahimian, R., Cordeau, P., & Kriz, J. (2019, 2019/05/01/). Brain Response to Injuries: When Microglia Go Sexist. *Neuroscience, 405*, 14-23. https://doi.org/https://doi.org/10.1016/j.neuroscience.2018.02.048

Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A., & Tyagi, S. (2008, Oct). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods, 5*(10), 877-879. https://doi.org/10.1038/nmeth.1253

Ransohoff, R. M. (2016, 2016/08/01). A polarizing question: do M1 and M2 microglia exist? *Nature Neuroscience, 19*(8), 987-991. https://doi.org/10.1038/nn.4338

Ransohoff, R. M., & Cardona, A. E. (2010, 2010/11/01). The myeloid cells of the central nervous system parenchyma. *Nature, 468*(7321), 253-262. https://doi.org/10.1038/nature09615

Reboldi, A., Coisne, C., Baumjohann, D., Benvenuto, F., Bottinelli, D., Lira, S., Uccelli, A., Lanzavecchia, A., Engelhardt, B., & Sallusto, F. (2009, 2009/05/01). C-C chemokine receptor 6–regulated entry of TH-17 cells into the CNS through the choroid plexus is required for the initiation of EAE. *Nature Immunology, 10*(5), 514-523. https://doi.org/10.1038/ni.1716

Reichard, A., & Asosingh, K. (2019, 2019/02/01). Best Practices for Preparing a Single Cell Suspension from Solid Tissues for Flow Cytometry [https://doi.org/10.1002/cyto.a.23690]. *Cytometry Part A, 95*(2), 219-226. https://doi.org/https://doi.org/10.1002/cyto.a.23690

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., Wadi, L., Meyer, M., Wong, J., Xu, C., Merico, D., & Bader, G. D. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols, 14*(2), 482-517. https://doi.org/10.1038/s41596-018-0103-9

Reimegård, J., Tarbier, M., Danielsson, M., Schuster, J., Baskaran, S., Panagiotou, S., Dahl, N., Friedländer, M. R., & Gallant, C. J. (2021, 2021/05/25). A combined approach for single-cell mRNA and intracellular protein expression analysis. *Communications Biology, 4*(1), 624. https://doi.org/10.1038/s42003-021-02142-w

Reu, P., Khosravi, A., Bernard, S., Mold, J. E., Salehpour, M., Alkass, K., Perl, S., Tisdale, J., Possnert, G., Druid, H., & Frisen, J. (2017, Jul 25). The Lifespan and Turnover of

Microglia in the Human Brain. *Cell Rep, 20*(4), 779-784. https://doi.org/10.1016/j.celrep.2017.07.004

Roberts, K., Aivazidis, A., Kleshchevnikov, V., Li, T., Fropf, R., Rhodes, M., Beechem, J. M., Hemberg, M., & Bayraktar, O. A. (2021). Transcriptome-wide spatial RNA profiling maps the cellular architecture of the developing human neocortex. *bioRxiv*, 2021.2003.2020.436265. https://doi.org/10.1101/2021.03.20.436265

Rowland, J. W., Hawryluk, G. W., Kwon, B., & Fehlings, M. G. (2008). Current status of acute spinal cord injury pathophysiology and emerging therapies: promise on the horizon. *Neurosurg Focus, 25*(5), E2. https://doi.org/10.3171/foc.2008.25.11.E2

Rudi, K., Zimonja, M., Trosvik, P., & Næs, T. (2007, 2007/11/30/). Use of multivariate statistics for 16S rRNA gene analysis of microbial communities. *International Journal of Food Microbiology, 120*(1), 95-99. https://doi.org/https://doi.org/10.1016/j.ijfoodmicro.2007.06.004

Rust, R., & Kaiser, J. (2017). Insights into the Dual Role of Inflammation after Spinal Cord Injury. *The Journal of Neuroscience, 37*(18), 4658. https://doi.org/10.1523/JNEUROSCI.0498-17.2017

Sanz, E., Yang, L., Su, T., Morris, D. R., McKnight, G. S., & Amieux, P. S. (2009). Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proceedings of the National Academy of Sciences, 106*(33), 13939. https://doi.org/10.1073/pnas.0907143106

Scheff, S., & Roberts, K. N. (2009). Infinite Horizon Spinal Cord Contusion Model. In J. Chen, Z. C. Xu, X.-M. Xu, & J. H. Zhang (Eds.), *Animal Models of Acute Neurological Injuries* (pp. 423-432). Humana Press. https://doi.org/10.1007/978-1-60327-185-1_36

Schulz, C., Perdiguero, E. G., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S. E. W., Pollard, J. W., Frampton, J., Liu, K. J., & Geissmann, F. (2012). A Lineage of Myeloid Cells Independent of Myb and Hematopoietic Stem Cells. *Science, 336*(6077), 86. https://doi.org/10.1126/science.1219179

Schwab, J. M., Zhang, Y., Kopp, M. A., Brommer, B., & Popovich, P. G. (2014, Aug). The paradox of chronic neuroinflammation, systemic immune suppression, autoimmunity after traumatic chronic spinal cord injury. *Exp Neurol, 258*, 121-129. https://doi.org/10.1016/j.expneurol.2014.04.023

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011, 2011/05/01). Global quantification of mammalian gene expression control. *Nature, 473*(7347), 337-342. https://doi.org/10.1038/nature10098

Scialdone, A., Natarajan, K. N., Saraiva, L. R., Proserpio, V., Teichmann, S. A., Stegle, O., Marioni, J. C., & Buettner, F. (2015, Sep 1). Computational assignment of cell-cycle

stage from single-cell transcriptome data. *Methods, 85*, 54-61. https://doi.org/10.1016/j.ymeth.2015.06.021

Seifert, J. L., Bell, J. E., Elmer, B. B., Sucato, D. J., & Romero, M. I. (2011, 2011/04/15/). Characterization of a novel bidirectional distraction spinal cord injury animal model. *Journal of Neuroscience Methods, 197*(1), 97-103. https://doi.org/https://doi.org/10.1016/j.jneumeth.2011.02.003

Seitz, A., Kragol, M., Aglow, E., Showe, L., & Heber-Katz, E. (2003, Feb 1). Apolipoprotein E expression after spinal cord injury in the mouse. *J Neurosci Res, 71*(3), 417-426. https://doi.org/10.1002/jnr.10482

Senga, S., Kobayashi, N., Kawaguchi, K., Ando, A., & Fujii, H. (2018, Sep). Fatty acid-binding protein 5 (FABP5) promotes lipolysis of lipid droplets, de novo fatty acid (FA) synthesis and activation of nuclear factor-kappa B (NF-κB) signaling in cancer cells. *Biochim Biophys Acta Mol Cell Biol Lipids, 1863*(9), 1057-1067. https://doi.org/10.1016/j.bbalip.2018.06.010

Seo, J., Jeong, D.-W., Park, J.-W., Lee, K.-W., Fukuda, J., & Chun, Y.-S. (2020, 2020/10/30). Fatty-acid-induced FABP5/HIF-1 reprograms lipid metabolism and enhances the proliferation of liver cancer cells. *Communications Biology, 3*(1), 638. https://doi.org/10.1038/s42003-020-01367-5

Seok, J., Warren, H. S., Cuenca, A. G., Mindrinos, M. N., Baker, H. V., Xu, W., Richards, D. R., McDonald-Smith, G. P., Gao, H., Hennessy, L., Finnerty, C. C., López, C. M., Honari, S., Moore, E. E., Minei, J. P., Cuschieri, J., Bankey, P. E., Johnson, J. L., Sperry, J., Nathens, A. B., Billiar, T. R., West, M. A., Jeschke, M. G., Klein, M. B., Gamelli, R. L., Gibran, N. S., Brownstein, B. H., Miller-Graziano, C., Calvano, S. E., Mason, P. H., Cobb, J. P., Rahme, L. G., Lowry, S. F., Maier, R. V., Moldawer, L. L., Herndon, D. N., Davis, R. W., Xiao, W., Tompkins, R. G., Inflammation, & Host Response to Injury, L. S. C. R. P. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences of the United States of America, 110*(9), 3507-3512. https://doi.org/10.1073/pnas.1222878110

Sharif-Alhoseini, M., Khormali, M., Rezaei, M., Safdarian, M., Hajighadery, A., Khalatbari, M. M., Safdarian, M., Meknatkhah, S., Rezvan, M., Chalangari, M., Derakhshan, P., & Rahimi-Movaghar, V. (2017, 2017/08/01). Animal models of spinal cord injury: a systematic review. *Spinal Cord, 55*(8), 714-721. https://doi.org/10.1038/sc.2016.187

Sheila-10X. (2017). *Single-Cell RNA-Seq: An Introductory Overview and Tools for Getting Started*. 10X Genomics. https://www.10xgenomics.com/blog/single-cell-rna-seq-an-introductory-overview-and-tools-for-getting-started

Shi, L.-L., Zhang, N., Xie, X.-M., Chen, Y.-J., Wang, R., Shen, L., Zhou, J.-S., Hu, J.-G., & Lü, H.-Z. (2017, February 15). Transcriptome profile of rat genes in injured spinal cord at

different stages by RNA-sequencing [journal article]. *BMC Genomics, 18*(1), 173. https://doi.org/10.1186/s12864-017-3532-x

Shin, S., Walz, K. A., Archambault, A. S., Sim, J., Bollman, B. P., Koenigsknecht-Talboo, J., Cross, A. H., Holtzman, D. M., & Wu, G. F. (2014). Apolipoprotein E mediation of neuro-inflammation in a murine model of multiple sclerosis. *Journal of Neuroimmunology, 271*(1-2), 8-17. https://doi.org/10.1016/j.jneuroim.2014.03.010

Silva, N. A., Sousa, N., Reis, R. L., & Salgado, A. J. (2014, Mar). From basics to clinical: a comprehensive review on spinal cord injury. *Prog Neurobiol, 114*, 25-57. https://doi.org/10.1016/j.pneurobio.2013.11.002

Singh, A., Tetreault, L., Kalsi-Ryan, S., Nouri, A., & Fehlings, M. G. (2014, 09/23). Global prevalence and incidence of traumatic spinal cord injury. *Clinical Epidemiology, 6*, 309-331. https://doi.org/10.2147/CLEP.S68889

Sinha, R., Stanley, G., Gulati, G. S., Ezran, C., Travaglini, K. J., Wei, E., Chan, C. K. F., Nabhan, A. N., Su, T., Morganti, R. M., Conley, S. D., Chaib, H., Red-Horse, K., Longaker, M. T., Snyder, M. P., Krasnow, M. A., & Weissman, I. L. (2017). Index switching causes "spreading-of-signal" among multiplexed samples in Illumina HiSeq 4000 DNA sequencing. *bioRxiv*, 125724. https://doi.org/10.1101/125724

Skene, P. J., & Henikoff, S. (2017, 2017/01/12). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife, 6*, e21856. https://doi.org/10.7554/eLife.21856

Smith, J., Braga, A., Hamel, R., Rutigliani, C., Yu, B., Peruzzotti-Jametti, L., & Pluchino, S. (2020). The therapeutic potential of exogenous adult stem cells for the injured central nervous system. In *Handbook of Innovations in Central Nervous System Regenerative Medicine*
(1st ed., pp. 147-258). https://doi.org/10.1016/b978-0-12-818084-6.00006-4

Sofroniew, M. V. (2015). Astrogliosis. *Cold Spring Harbor Perspectives in Biology, 7*(2), a020420. https://doi.org/10.1101/cshperspect.a020420

Spiering, M. J. (2015). Primer on the Immune System. *Alcohol research : current reviews, 37*(2), 171-175. https://pubmed.ncbi.nlm.nih.gov/26695756
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4590614/

Sroga, J. M., Jones, T. B., Kigerl, K. A., McGaughy, V. M., & Popovich, P. G. (2003). Rats and mice exhibit distinct inflammatory reactions after spinal cord injury. *Journal of Comparative Neurology, 462*(2), 223-240. https://doi.org/10.1002/cne.10736

Stirling, D. P., Liu, S., Kubes, P., & Yong, V. W. (2009, Jan 21). Depletion of Ly6G/Gr-1 leukocytes after spinal cord injury in mice alters wound healing and worsens neurological outcome. *The Journal of neuroscience : the official journal of the Society*

for Neuroscience, 29(3), 753-764. https://doi.org/10.1523/jneurosci.4918-08.2009

Stirling, D. P., & Yong, V. W. (2008, Jul). Dynamics of the inflammatory response after murine spinal cord injury revealed by flow cytometry. *J Neurosci Res, 86*(9), 1944-1958. https://doi.org/10.1002/jnr.21659

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P. K., Swerdlow, H., Satija, R., & Smibert, P. (2017, Sep). Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods, 14*(9), 865-868. https://doi.org/10.1038/nmeth.4380

Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B. Z., Mauck, W. M., Smibert, P., & Satija, R. (2018, 2018/12/19). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biology, 19*(1), 224. https://doi.org/10.1186/s13059-018-1603-1

Sun, G., Yang, S., Cao, G., Wang, Q., Hao, J., Wen, Q., Li, Z., So, K.-F., Liu, Z., Zhou, S., Zhao, Y., Yang, H., Zhou, L., & Yin, Z. (2018). γδ T cells provide the early source of IFN-γ to aggravate lesions in spinal cord injury. *The Journal of experimental medicine, 215*(2), 521-535. https://doi.org/10.1084/jem.20170686

Swirski, F. K., Hilgendorf, I., & Robbins, C. S. (2014). From proliferation to proliferation: monocyte lineage comes full circle. *Seminars in immunopathology, 36*(2), 137-148. https://doi.org/10.1007/s00281-013-0409-1

Tasic, B., Menon, V., Nguyen, T. N., Kim, T. K., Jarsky, T., Yao, Z., Levi, B., Gray, L. T., Sorensen, S. A., Dolbeare, T., Bertagnolli, D., Goldy, J., Shapovalova, N., Parry, S., Lee, C., Smith, K., Bernard, A., Madisen, L., Sunkin, S. M., Hawrylycz, M., Koch, C., & Zeng, H. (2016, 2016/02/01). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature Neuroscience, 19*(2), 335-346. https://doi.org/10.1038/nn.4216

Tay, T. L., Mai, D., Dautzenberg, J., Fernández-Klett, F., Lin, G., Sagar, Datta, M., Drougard, A., Stempfl, T., Ardura-Fabregat, A., Staszewski, O., Margineanu, A., Sporbert, A., Steinmetz, L. M., Pospisilik, J. A., Jung, S., Priller, J., Grün, D., Ronneberger, O., & Prinz, M. (2017, 2017/06/01). A new fate mapping system reveals context-dependent random or clonal expansion of microglia. *Nature Neuroscience, 20*(6), 793-803. https://doi.org/10.1038/nn.4547

Tay, T. L., Sagar, Dautzenberg, J., Grün, D., & Prinz, M. (2018, 2018/09/05). Unique microglia recovery population revealed by single-cell RNAseq following neurodegeneration. *Acta Neuropathologica Communications, 6*(1), 87. https://doi.org/10.1186/s40478-018-0584-3

Thrupp, N., Sala Frigerio, C., Wolfs, L., Skene, N. G., Fattorelli, N., Poovathingal, S., Fourne, Y., Matthews, P. M., Theys, T., Mancuso, R., de Strooper, B., & Fiers, M. (2020,

2020/09/29/). Single-Nucleus RNA-Seq Is Not Suitable for Detection of Microglial Activation Genes in Humans. *Cell Rep, 32*(13), 108189. https://doi.org/https://doi.org/10.1016/j.celrep.2020.108189

Traag, V. A., Waltman, L., & van Eck, N. J. (2019, 2019/03/26). From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports, 9*(1), 5233. https://doi.org/10.1038/s41598-019-41695-z

Tran, H. T. N., Ang, K. S., Chevrier, M., Zhang, X., Lee, N. Y. S., Goh, M., & Chen, J. (2020, 2020/01/16). A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biology, 21*(1), 12. https://doi.org/10.1186/s13059-019-1850-9

Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome research, 25*(10), 1491-1498. https://doi.org/10.1101/gr.190595.115

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S., & Rinn, J. L. (2014, 2014/04/01). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol, 32*(4), 381-386. https://doi.org/10.1038/nbt.2859

Trivedi, A., Olivas, A. D., & Noble-Haeusslein, L. J. (2006, 2006/12//). Inflammation and Spinal Cord Injury: Infiltrating Leukocytes as Determinants of Injury and Repair Processes. *Clinical neuroscience research, 6*(5), 283-292. https://doi.org/10.1016/j.cnr.2006.09.007

Tyler, J. Y., Xu, X. M., & Cheng, J. X. (2013, Oct 7). Nanomedicine for treating spinal cord injury. *Nanoscale, 5*(19), 8821-8836. https://doi.org/10.1039/c3nr00957b

van den Brink, S. C., Sage, F., Vértesy, Á., Spanjaard, B., Peterson-Maduro, J., Baron, C. S., Robin, C., & van Oudenaarden, A. (2017, 2017/10/01). Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature Methods, 14*(10), 935-936. https://doi.org/10.1038/nmeth.4437

van der Maaten, L. (2013). *Visualizing Data Using t-SNE* [Presentation]. https://www.youtube.com/watch?v=RJVL80Gg3lA

van der Maaten, L. H., G. (2008). Visualizing High-Dimensional Data Using t-SNE. *Journal of Machine Learning Research, 9*, 2579-2605.

Van Hove, H., Martens, L., Scheyltjens, I., De Vlaminck, K., Pombo Antunes, A. R., De Prijck, S., Vandamme, N., De Schepper, S., Van Isterdael, G., Scott, C. L., Aerts, J., Berx, G., Boeckxstaens, G. E., Vandenbroucke, R. E., Vereecke, L., Moechars, D., Guilliams, M., Van Ginderachter, J. A., Saeys, Y., & Movahedi, K. (2019, 2019/06/01). A single-cell atlas of mouse brain macrophages reveals unique transcriptional identities shaped by

ontogeny and tissue environment. *Nature Neuroscience, 22*(6), 1021-1035. https://doi.org/10.1038/s41593-019-0393-4

Vaughn, C. N., Iafrate, J. L., Henley, J. B., Stevenson, E. K., Shlifer, I. G., & Jones, T. B. (2013, 2013/08/01). Cellular Neuroinflammation in a Lateral Forceps Compression Model of Spinal Cord Injury [https://doi.org/10.1002/ar.22730]. *The Anatomical Record, 296*(8), 1229-1246. https://doi.org/https://doi.org/10.1002/ar.22730

Vento-Tormo, R., Efremova, M., Botting, R. A., Turco, M. Y., Vento-Tormo, M., Meyer, K. B., Park, J.-E., Stephenson, E., Polański, K., Goncalves, A., Gardner, L., Holmqvist, S., Henriksson, J., Zou, A., Sharkey, A. M., Millar, B., Innes, B., Wood, L., Wilbrey-Clark, A., Payne, R. P., Ivarsson, M. A., Lisgo, S., Filby, A., Rowitch, D. H., Bulmer, J. N., Wright, G. J., Stubbington, M. J. T., Haniffa, M., Moffett, A., & Teichmann, S. A. (2018, 2018/11/01). Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature, 563*(7731), 347-353. https://doi.org/10.1038/s41586-018-0698-6

Vucic, D., Dixit, V. M., & Wertz, I. E. (2011, Jun 23). Ubiquitylation in apoptosis: a post-translational modification at the edge of life and death. *Nat Rev Mol Cell Biol, 12*(7), 439-452. https://doi.org/10.1038/nrm3143

Wahane, S., Zhou, X., Zhou, X., Guo, L., Friedl, M.-S., Kluge, M., Ramakrishnan, A., Shen, L., Friedel, C. C., Zhang, B., Friedel, R. H., & Zou, H. (2021). Diversified transcriptional responses of myeloid and glial cells in spinal cord injury shaped by HDAC3 activity. *Science Advances, 7*(9), eabd8811. https://doi.org/10.1126/sciadv.abd8811

Wang, J., Gan, Y., Han, P., Yin, J., Liu, Q., Ghanian, S., Gao, F., Gong, G., & Tang, Z. (2018, 2018/01/11). Ischemia-induced Neuronal Cell Death Is Mediated by Chemokine Receptor CX3CR1. *Scientific Reports, 8*(1), 556. https://doi.org/10.1038/s41598-017-18774-0

Wang, X., Cao, K., Sun, X., Chen, Y., Duan, Z., Sun, L., Guo, L., Bai, P., Sun, D., Fan, J., He, X., Young, W., & Ren, Y. (2015, 2015/04/01). Macrophages in spinal cord injury: Phenotypic and functional change from exposure to myelin debris. *Glia, 63*(4), 635-651. https://doi.org/10.1002/glia.22774

Wang, Y., Cella, M., Mallinson, K., Ulrich, Jason D., Young, Katherine L., Robinette, Michelle L., Gilfillan, S., Krishnan, Gokul M., Sudhakar, S., Zinselmeyer, Bernd H., Holtzman, David M., Cirrito, John R., & Colonna, M. (2015, 2015/03/12/). TREM2 Lipid Sensing Sustains the Microglial Response in an Alzheimer's Disease Model. *Cell, 160*(6), 1061-1071. https://doi.org/https://doi.org/10.1016/j.cell.2015.01.049

Warburg, O. (1956, Feb 24). On the origin of cancer cells. *Science, 123*(3191), 309-314. https://doi.org/10.1126/science.123.3191.309

Whetstone, W. D., Hsu, J.-Y. C., Eisenberg, M., Werb, Z., & Noble-Haeusslein, L. J. (2003, 2003/10/15). Blood-spinal cord barrier after spinal cord injury: Relation to

revascularization and wound healing [https://doi.org/10.1002/jnr.10759]. *Journal of Neuroscience Research, 74*(2), 227-239. https://doi.org/https://doi.org/10.1002/jnr.10759

Willemsen, L., & de Winther, M. P. (2020). Macrophage subsets in atherosclerosis as defined by single-cell technologies. *The Journal of pathology, 250*(5), 705-714. https://doi.org/10.1002/path.5392

Witiw, C. D., & Fehlings, M. G. (2015, Jul). Acute Spinal Cord Injury. *J Spinal Disord Tech, 28*(6), 202-210. https://doi.org/10.1097/bsd.0000000000000287

Wu, T., Tian, J., Cutler, R. G., Telljohann, R. S., Bernlohr, D. A., Mattson, M. P., & Handa, J. T. (2010). Knockdown of FABP5 mRNA decreases cellular cholesterol levels and results in decreased apoB100 secretion and triglyceride accumulation in ARPE-19 cells. *Laboratory investigation; a journal of technical methods and pathology, 90*(6), 906-914. https://doi.org/10.1038/labinvest.2009.33

Xu, L., Botchway, B. O. A., Zhang, S., Zhou, J., & Liu, X. (2018). Inhibition of NF-κB Signaling Pathway by Resveratrol Improves Spinal Cord Injury. *Frontiers in neuroscience, 12*, 690-690. https://doi.org/10.3389/fnins.2018.00690

Xu, P., Zhang, F., Chang, M.-m., Zhong, C., Sun, C.-H., Zhu, H.-R., Yao, J.-C., Li, Z.-Z., Li, S.-T., Zhang, W.-C., & Sun, G.-D. (2021, 2021/03/02). Recruitment of γδ T cells to the lesion via the CCL2/CCR2 signaling after spinal cord injury. *J Neuroinflammation, 18*(1), 64. https://doi.org/10.1186/s12974-021-02115-0

Yaguchi, M., Tabuse, M., Ohta, S., Ohkusu-Tsukada, K., Takeuchi, T., Yamane, J., Katoh, H., Nakamura, M., Matsuzaki, Y., Yamada, M., Itoh, T., Nomura, T., Toyama, Y., Okano, H., & Toda, M. (2009, 2009/12/01/). Transplantation of dendritic cells promotes functional recovery from spinal cord injury in common marmoset. *Neuroscience Research, 65*(4), 384-392. https://doi.org/https://doi.org/10.1016/j.neures.2009.08.016

Yamawaki, T. M., Lu, D. R., Ellwanger, D. C., Bhatt, D., Manzanillo, P., Arias, V., Zhou, H., Yoon, O. K., Homann, O., Wang, S., & Li, C.-M. (2021, 2021/01/20). Systematic comparison of high-throughput single-cell RNA-seq methods for immune cell profiling. *BMC Genomics, 22*(1), 66. https://doi.org/10.1186/s12864-020-07358-4

Yang, T., Dai, Y., Chen, G., & Cui, S. (2020, 2020-April-03). Dissecting the Dual Role of the Glial Scar and Scar-Forming Astrocytes in Spinal Cord Injury [Review]. *Frontiers in cellular neuroscience, 14*(78). https://doi.org/10.3389/fncel.2020.00078

Yang, X., Chen, S., Shao, Z., Li, Y., Wu, H., Li, X., Mao, L., Zhou, Z., Bai, L., Mei, X., & Liu, C. (2018). Apolipoprotein E Deficiency Exacerbates Spinal Cord Injury in Mice: Inflammatory Response and Oxidative Stress Mediated by NF-κB Signaling Pathway. *Frontiers in cellular neuroscience, 12*, 142-142. https://doi.org/10.3389/fncel.2018.00142

Yang, Y., Xin, X., Fu, X., & Xu, D. (2018, 2018/04/01). Expression pattern of human SERPINE2 in a variety of human tumors. *Oncol Lett, 15*(4), 4523-4530. https://doi.org/10.3892/ol.2018.7819

Yao, A.-H., Jia, L.-Y., Zhang, Y.-K., Ma, Q.-R., Cheng, P., Liu, L., Ju, G., & Kuang, F. (2012). Early Blockade of TLRs MyD88-Dependent Pathway May Reduce Secondary Spinal Cord Injury in the Rats. *Evidence-based complementary and alternative medicine : eCAM, 2012*, 591298-591298. https://doi.org/10.1155/2012/591298

Yoshizaki, S., Tamaru, T., Hara, M., Kijima, K., Tanaka, M., Konno, D.-j., Matsumoto, Y., Nakashima, Y., & Okada, S. (2021, 2021/01/06). Microglial inflammation after chronic spinal cord injury is enhanced by reactive astrocytes via the fibronectin/β1 integrin pathway. *J Neuroinflammation, 18*(1), 12. https://doi.org/10.1186/s12974-020-02059-x

Young, M. D., & Behjati, S. (2018). SoupX removes ambient RNA contamination from droplet based single cell RNA sequencing data. *bioRxiv*, 303727. https://doi.org/10.1101/303727

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010, 2010/02/04). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology, 11*(2), R14. https://doi.org/10.1186/gb-2010-11-2-r14

Zeisel, A., Hochgerner, H., Lonnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Haring, M., Braun, E., Borm, L. E., La Manno, G., Codeluppi, S., Furlan, A., Lee, K., Skene, N., Harris, K. D., Hjerling-Leffler, J., Arenas, E., Ernfors, P., Marklund, U., & Linnarsson, S. (2018, Aug 9). Molecular Architecture of the Mouse Nervous System. *Cell, 174*(4), 999-1014.e1022. https://doi.org/10.1016/j.cell.2018.06.021

Zhang, A. W., O'Flanagan, C., Chavez, E. A., Lim, J. L. P., Ceglia, N., McPherson, A., Wiens, M., Walters, P., Chan, T., Hewitson, B., Lai, D., Mottok, A., Sarkozy, C., Chong, L., Aoki, T., Wang, X., Weng, A. P., McAlpine, J. N., Aparicio, S., Steidl, C., Campbell, K. R., & Shah, S. P. (2019, 2019/10/01). Probabilistic cell-type assignment of single-cell RNA-seq for tumor microenvironment profiling. *Nature Methods, 16*(10), 1007-1015. https://doi.org/10.1038/s41592-019-0529-1

Zhang, B., Zou, J., Han, L., Beeler, B., Friedman, J. L., Griffin, E., Piao, Y.-S., Rensing, N. R., & Wong, M. (2018). The specificity and role of microglia in epileptogenesis in mouse models of tuberous sclerosis complex. *Epilepsia, 59*(9), 1796-1806. https://doi.org/10.1111/epi.14526

Zhang, C., Morozova, A. Y., Abakumov, M. A., Gubsky, I. L., Douglas, P., Feng, S., Bryukhovetskiy, A. S., & Chekhonin, V. P. (2015). Precise Delivery Into Chronic Spinal Cord Injury Syringomyelic Cysts with Magnetic Nanoparticles MRI Visualization. *Medical science monitor : international medical journal of experimental and clinical research, 21*, 3179-3185. https://doi.org/10.12659/msm.895624

Zhang, N., Fang, M., Chen, H., Gou, F., & Ding, M. (2014, 09/08/accepted). Evaluation of spinal cord injury animal models. *Neural Regeneration Research, 9*(22), 2008-2012. https://doi.org/10.4103/1673-5374.143436

Zhang, Y., Sun, Y., Rao, E., Yan, F., Li, Q., Zhang, Y., Silverstein, K. A., Liu, S., Sauter, E., Cleary, M. P., & Li, B. (2014, Jun 1). Fatty acid-binding protein E-FABP restricts tumor growth by promoting IFN-β responses in tumor-associated macrophages. *Cancer Res, 74*(11), 2986-2998. https://doi.org/10.1158/0008-5472.Can-13-2689

Zhang, Y. L., Hernandez-Ono, A., Ko, C., Yasunaga, K., Huang, L. S., & Ginsberg, H. N. (2004, Apr 30). Regulation of hepatic apolipoprotein B-lipoprotein assembly and secretion by the availability of fatty acids. I. Differential response to the delivery of fatty acids via albumin or remnant-like emulsion particles. *The Journal of biological chemistry, 279*(18), 19362-19374. https://doi.org/10.1074/jbc.M400220200

Zhao, G., Wu, M., Wang, X., Du, Z., & Zhang, G. (2017, 2017/10/01). Effect of FABP5 gene silencing on the proliferation, apoptosis and invasion of human gastric SGC-7901 cancer cells. *Oncol Lett, 14*(4), 4772-4778. https://doi.org/10.3892/ol.2017.6748

Zhao, X.-F., Alam, M. M., Liao, Y., Huang, T., Mathur, R., Zhu, X., & Huang, Y. (2019). Targeting Microglia Using Cx3cr1-Cre Lines: Revisiting the Specificity. *eNeuro, 6*(4), ENEURO.0114-0119.2019. https://doi.org/10.1523/ENEURO.0114-19.2019

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., Bharadwaj, R., Wong, A., Ness, K. D., Beppu, L. W., Deeg, H. J., McFarland, C., Loeb, K. R., Valente, W. J., Ericson, N. G., Stevens, E. A., Radich, J. P., Mikkelsen, T. S., Hindson, B. J., & Bielas, J. H. (2017, 01/16/online). Massively parallel digital transcriptional profiling of single cells [Article]. *Nat Commun, 8*, 14049. https://doi.org/10.1038/ncomms14049
https://www.nature.com/articles/ncomms14049#supplementary-information

Zhou, X., He, X., & Ren, Y. (2014, 08/20/accepted). Function of microglia and macrophages in secondary damage after spinal cord injury. *Neural Regeneration Research, 9*(20), 1787-1795. https://doi.org/10.4103/1673-5374.143423

Zhou, X., Wahane, S., Friedl, M. S., Kluge, M., Friedel, C. C., Avrampou, K., Zachariou, V., Guo, L., Zhang, B., He, X., Friedel, R. H., & Zou, H. (2020, Mar). Microglia and macrophages promote corralling, wound compaction and recovery after spinal cord injury via Plexin-B2. *Nat Neurosci, 23*(3), 337-350. https://doi.org/10.1038/s41593-020-0597-7

Zhu, Y., Lyapichev, K., Lee, D. H., Motti, D., Ferraro, N. M., Zhang, Y., Yahn, S., Soderblom, C., Zha, J., Bethea, J. R., Spiller, K. L., Lemmon, V. P., & Lee, J. K. (2017, Mar 1). Macrophage Transcriptional Profile Identifies Lipid Catabolic Pathways That Can Be Therapeutically Targeted after Spinal Cord Injury. *The Journal of neuroscience : the*

official journal of the Society for Neuroscience, 37(9), 2362-2376. https://doi.org/10.1523/jneurosci.2751-16.2017

Zhu, Y., Soderblom, C., Krishnan, V., Ashbaugh, J., Bethea, J. R., & Lee, J. K. (2015, 2015/02/01/). Hematogenous macrophage depletion reduces the fibrotic scar and increases axonal growth after spinal cord injury. *Neurobiology of Disease, 74*, 114-125. https://doi.org/https://doi.org/10.1016/j.nbd.2014.10.024

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., & Enard, W. (2017, 2017/02/16/). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell, 65*(4), 631-643.e634. https://doi.org/https://doi.org/10.1016/j.molcel.2017.01.023

Zivkovic, S., Ayazi, M., Hammel, G., & Ren, Y. (2021, 2021-April-22). For Better or for Worse: A Look Into Neutrophils in Traumatic Spinal Cord Injury [Review]. *Frontiers in cellular neuroscience, 15*(91). https://doi.org/10.3389/fncel.2021.648076