Chapter 1: Introduction

1.1 Identification and characterisation of microglial cells in the brain

Microglia are the tissue resident macrophages of the central nervous system (CNS) and play an important role in it's immune defense[1]. Microglia were first described in the early 1900s, as scientists began to use developing microscopy techniques to study the brain. Santiago Ramón y Cajal, a Spanish neuroscientist famed for his descriptions and images of the CNS, dedicated much of his research to the non-neuronal cells within the brain, known as glial cells[2,3]. Within this glial cell population, Cajal identified the "third element" of the CNS describing the non-neuronal, non-astrocytic population of cells he observed. Río-Hortega divided this "third element" into two subdivisions: microglia and interfascicular glia, now known as oligodendrocytes[2]. Río-Hortega observed that microglia were relatively uniformly distributed in the brain, although noted a higher density in the grey matter, and described the cells as highly dynamic, often adapting their morphology to the features of the brain[3]. His later work focussed on microglial physiology following trauma to the brain where he described the cells taking on an ameboid shape and becoming highly phagocytic.

Since the early description of microglia, experimental tools have significantly improved and it is now easier to identify and observe microglial cells in a variety of systems, from primary cells across species to *in-vitro* models. Improved microscopy techniques have confirmed Rio-Hortega's initial observation that microglia have a highly ramified morphology (Figure 1.1), with dynamic processes that constantly survey the environment and maintain contact with neurons[4]. *In-vivo* time lapse imaging using zebrafish has suggested that this motility is not a random process[5] and that the cells are responding to ATP signals released from active neurons.

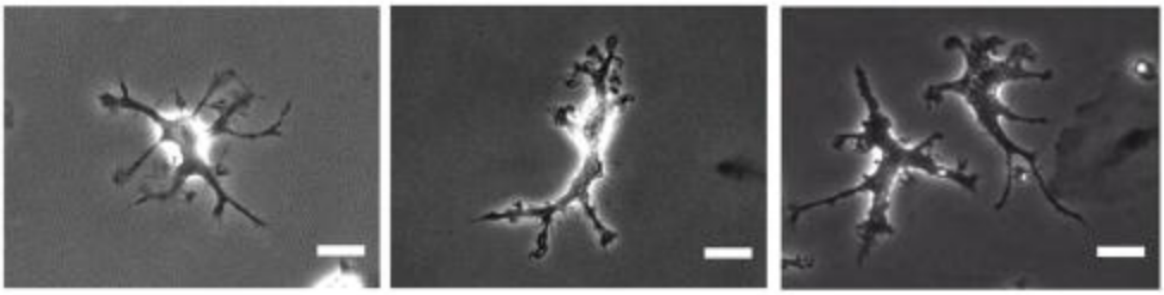

Figure 1.1 Microscopy images of mouse (left), fetal human (middle) and iPSC-derived microglia

Image taken from Muffat *et al*.[6], Figure 3 panel b.

In addition to describing the characteristics of microglial cells, Rio-Hortega was the first to theorise that microglia were of mesoderm origin[7]. For many years this theory was overlooked and instead it was argued that the cells were derived from neuro-ectoderm, along with other glial cell populations such as astrocytes[8–10]. However, evidence began to build that supported Rio-Hortega's original proposal: microglia were shown to have similar morphological features to macrophages[11] and were shown to express myeloid markers such as CD11b[12]. In mice, knockout (KO) of the *PU.1* gene, a key transcription factor (TF) in myeloid cell development, resulted in an absence of microglial populations in the brain[13].

1.2 Lineage of microglial populations in the brain

It is now well recognised that the microglial cells first described by Río-Hortega are tissue resident macrophages of the CNS. While the myeloid origin of these cells is no longer disputed, unique features of microglial development appear to distinguish them from other macrophage cells both in their initial origin and maintenance throughout adult life.

1.2.1 Microglial cell origin in embryonic development

Microglia-like cells have been identified in both rodent and human samples in the very early stages of embryonic development[14,15], suggesting they derive from a lineage independent of bone marrow hematopoiesis. In human fetal development,

Iba1+ (a myeloid cell marker) precursor cells have been observed in the developing nervous system as early as 4.5 gestational weeks[15], while hematopoietic stem cells don't seed the fetal liver until around gestational week 5[16].

Dissociation of fetal tissue samples from mice provided the first evidence that microglial progenitors are located in the yolk sac (YS) before moving into the developing brain as embryogenesis progresses[14]. More recently, a fate-mapping study has provided further evidence of the unique YS origin of microglial cells[17]. Fate-mapping relies on the ability to label cells from specific developmental origins and trace them through the developmental process. In the case of microglia, yellow fluorescent labelled protein (YFP) was linked to the *RUNX1* TF, which is specific to YS myeloid development. An estimated 32% of adult microglia cells were derived from YS precursors compared to only 3% of circulating monocytes. Specific erythro-myeloid progenitors within the mouse YS have since been identified[18] and it is these colony stimulating factor 1 receptor (CSF-1R) expressing-cells that appear to give rise to tissue resident macrophages such as microglia.

Mouse models have also been used to identify the pathways and molecules that regulate microglial differentiation from early progenitors. *Myb* is a TF which has previously been shown to be dispensable for yolk sac myelopoiesis but necessary for the creation of hematopoietic stem cells in the bone marrow. The initial production of microglia cells has been shown to be a *Myb* independent process[19,20], which further adds to the evidence behind the YS origin of microglia. Other TFs, like *PU.1* and *IRF8*, as well as protein coding genes, such as *MMP8* and *MMP9*, are required for the development of mature microglial cells[19,21]. The expression of CSF-1R by progenitor cells and a functional circulatory system is also necessary for microglial differentiation[17].

1.2.2 Maintenance of microglial populations throughout adulthood

The CNS has long been considered an "immune privileged" site, which limits immune reactions in the brain[22]. This, in part, is due to the presence of the blood brain barrier (BBB) that is thought to prevent circulating immune cells entering the brain. In most other tissues, circulating monocytes provide a progenitor cell for expanding

macrophage populations. It is known that even after the formation of the BBB, when monocytes theoretically cannot enter the brain, the population of microglia in the brain continues to grow with a large population surge two weeks after birth[14]. This evidence suggests that microglial cells have expansion potential and can self-maintain populations throughout adulthood. There are three proposed mechanisms for this continued growth of microglial populations: i) microglia are in fact replenished by circulating monocytes that cross the BBB, ii) there are populations of microglial progenitor cells that are present in the brain throughout life or iii) mature microglia themselves have the potential to proliferate.

Evidence for a significant contribution of circulating cells to the adult microglial population is controversial. Consistent with this hypothesis, *PU.1* KO mice lack any embryonically-derived microglia, but develop microglia-like cells within their CNS after receiving bone marrow transplants after birth[23]. However, fate-mapping studies have been used to demonstrate that up to 60% of microglia in adult mice are YS derived[20] and sublethal irradiation of mice followed by healthy hematopoietic cell transfer only gave rise to around 5% of donor derived microglia[17]. Parabiotic mouse models can be used to surgically join two mice and allow sharing of blood circulation, providing a useful tool for researchers to study how circulating cells contribute to certain populations[24]. If circulating monocytes contribute to the maintenance of homeostatic levels of microglia, one would expect to see similar levels of non-host cells in both the circulating system and the brain. However, multiple studies have demonstrated that parabiotic mice maintain higher levels of host-linked microglia[25–27] suggesting that monocyte cells do not contribute to the adult microglial population under normal conditions. It may be that under extreme conditions, such as a complete absence of microglia, brain injury or following significant neuroinflammation, circulating cells infiltrate the CNS. These cells may then contribute to the population of microglia-like cells in the brain, but this does not appear to be the case under homeostatic conditions[17,23,28].

The second theory of microglial repopulation is that there are progenitor cells within the brain that can differentiate into mature microglia. Following depletion of the microglial population in the adult mouse brain, using CSF-1R inhibitors, it has been

demonstrated that microglia rapidly repopulate the brain[29]. The rate of repopulation of microglia described in this study (from 600 cells/slice to >14,000 cells/slice in 72 hours) was determined to be too quick for repopulation to be explained by surviving cells. However, the presence of a progenitor population could explain these observations. Within the same study a population of Nestin and Ki67 positive cells were identified that appeared to be the source of repopulation. Initially, the nestin positive population had a distinct morphology to resident microglia, but then adopted the ramified morphology normally expected of native cells. However, since their initial description, the presence of microglia progenitor cells in the brain has remained controversial. Future studies have failed to identify a progenitor population[27] and noted that, while repopulating microglia may transiently express nestin, these cells derived solely from surviving cells. This suggests that adult microglia have proliferative potential and native cells are the driver behind population expansion.

1.3 Microglial function in development and the adult brain

There has been extensive research into the various roles microglia may play throughout the lifespan (Figure 1.2 [1,30–33]). As macrophage cells, microglia can clearly play an active role in the immune defense of the CNS. However, a growing body of evidence has shown that microglia are required for both neuronal development and normal brain function.

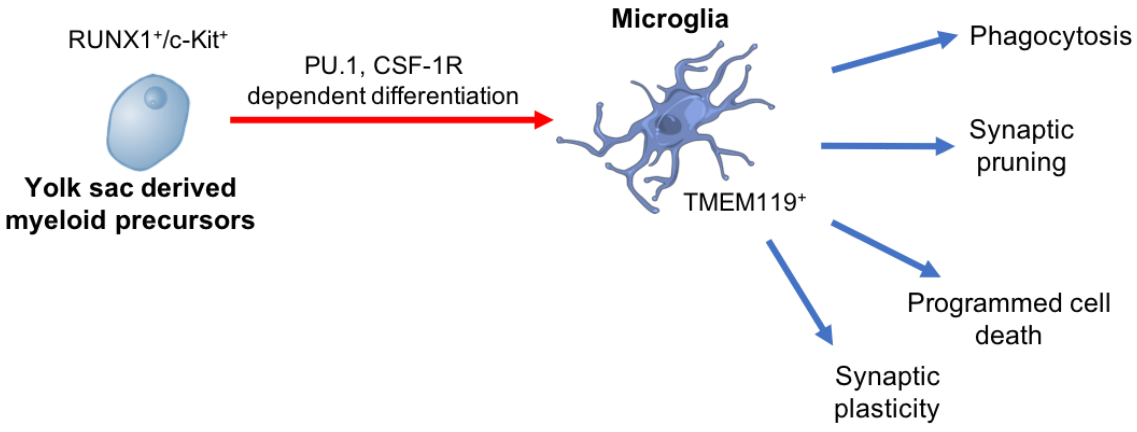

Figure 1.2 Overview of microglial development and function

A summary of microglial developmental pathways and functions in the healthy brain.

1.3.1 The role of microglia in the developing brain

Research in both humans and mice has demonstrated that microglia play an important role throughout brain development. Individuals with mutations in important regulators of microglial function, such as the *CSF-1R* gene, have profound neurological abnormalities[34] including abnormal arrangement of neurons and a lack of corpus callosum development. Studies like this provide direct evidence from human patients that microglial cells are required for normal brain development. However, these small scale patient studies cannot provide mechanistic details and so mouse models are often used as tools for studying microglia in development.

At the cellular level, microglia are able to phagocytose the early pool of neural precursor cells in order to control neurogenesis[35,36]. Studies have demonstrated that microglia play other important roles in brain development beyond their phagocytic function. Experimental evidence supports the idea that microglia provide trophic support to developing neurons in layer V cortical neurons in mice[37]. The cells accumulated close to the projection axons and, via a CX3CR1 dependent mechanism, produced IGF1 that maintained neuronal survival. Alongside trophic support for developing neurons microglial signalling has been shown to function in the programmed cell death of neurons. In the development of murine retina, prevention of microglial colonization of the tissue alleviated the production of nerve growth factor (NGF) and significantly reduced the level of normal programmed cell death[38]. More recent studies in both mouse Purkinje cells[39] and neurons in the mouse hippocampus[40] have implicated superoxide ions produced by microglia, through a CD11b/DAP12 dependent signalling pathway, in programmed cell death. Outside of their direct interactions with neurons, microglia also appear to be important for functional vasculature development *in-vivo* and, in the *in-vitro* based aortic ring model, addition of microglia cells to the culture stimulated vessel sprouting[41].

While the studies described above provide some evidence of the potential impact of microglia on neuronal development, one of the most well established and recognised functions of the cells in the developing brain is within the process of synaptic pruning. Synaptic pruning systematically removes weaker neurons and synaptic connections

to strengthen and improve the efficiency of the remaining connections within the brain. Experiments have shown that microglia closely co-localise to synaptic connections during active periods of pruning[42] and lysosomal markers have been used to highlight active engulfment of synaptic material[42,43]. Schafer *et al.*[42] studied microglia engulfment of synapses within mouse retinal ganglions and demonstrated that the cells preferentially digested "weaker" synaptic regions further supporting microglial involvement in the synaptic pruning process. Other studies have since established that the active engulfment of synapses by microglia is dependent on the activation of the classical complement cascade[42,44,45]. Disruptions of the CR3/C3 signaling cascade have been shown to cause deficits in synaptic connectivity[42] and *C1q* KO mice also have large disruptions in synapse elimination[44]. It is thought that complement protein tagged neurons provide the signal for phagocytosis by microglia[44].

1.3.2 Microglia in adulthood

Under normal conditions the brain is considered an "immune privileged" site, with the blood brain barrier (BBB) acting as a source of protection from infiltrating pathogens. While microglia may not have major immune functions under homeostatic conditions in the adult brain, it does not mean they remain inactive until disease or disruption occurs. Microglia are known to have a variety of homeostatic functions including phagocytosis of debris within the brain and monitoring of neuronal activity[1]. Many of the identified functions of microglial cells have been linked to CX3CR1 signalling. CX3CR1 is a receptor that is selectively expressed by microglia within the brain, which interacts with CX3CL1 ligand produced by neurons[46].

Recent evidence has also shown that microglia are important in the process of learning and memory in adults[47–49]. Learning and memory occur through the strengthening of synaptic and neuronal connections via processes of synaptic plasticity and long-term potentiation (LTP). *CX3CR1* KO mice have an impairment in measurable LTP alongside significant deficits in behavioural learning tests like fear conditioning and the Morris Water Maze[47]. ATP released by microglia in mice appears to modulate synaptic transmission by acting on $P2X_4$ and adenosine A1 receptors[48]. Using a selective eye closure mouse model, Sipe *et al.*[49] demonstrated

that microglia actively contribute to experience dependent plasticity through P2RY12 signalling.

There is also some evidence that external environmental factors can modulate microglial function. For instance, a high fat diet appeared to increase the number of microglia present within the hypothalamic region and was accompanied by an increased anti-inflammatory phenotype[50]. Obese humans studied within the same paper also showed cell type specific differences, including microglial dystrophy. Germ-free mice have also been used to study the impact of microbiome variation on microglial function[51]; without manipulation the mice showed global microglial defects including an immature phenotype and an impaired innate immune response. Recolonisation of germ-free mice partially restores microglia function, suggesting the influence of the gut microbiome on the brain is a dynamic process. However, these studies often do not provide evidence of specific molecular mechanisms that may drive these effects. Therefore, further research would need to be carried out to fully develop the scientific theories.

1.4 Microglia and disease

As the only major population of immune cells within the brain, microglia act as a first line of defence against infiltrating pathogens and are responsible for the clearance of cellular debris. However, microglia can also play a role in the development and progression of many disorders not immediately thought of as immune related [1,31,52]. When discussing microglia and disease it is important to distinguish between examples where microglia appear to play a causal role and those where the cells react to disease onset. The most well established causal link between microglial function and disease is Alzheimer's disease (AD) and as such this is discussed in more detail in section 1.5. The remainder of this section describes the evidence linking microglial function to a variety of other disorders and how the cells are involved in onset and progression.

1.4.1 Microglia in traumatic brain injury

Traumatic brain injury (TBI) is defined as “an alteration in brain function, or other evidence of brain pathology, caused by an external force”[53] and can often be further subdivided depending on the severity or outcome of the injury. As reactive immune cells within the brain, in the immediate aftermath of TBI microglial processes move rapidly to the site of injury, within minutes of damage[54]. Here, their primary function is to prevent disruption to the blood brain barrier[54–56]. Release of ATP from damaged tissue is thought to signal to microglia and stimulate the rapid movement of processes to the injury site, often without the movement of the cell body[54]. In mice it appears that microglial processes form specific honeycomb structures with single-process microglia dispersed throughout to assist with the sealing of the BBB[55]. A rapid increase in myeloid cell numbers occurs immediately in mice and can continue for up to four days[57]. Studies in human post-mortem brain samples have shown that the neuroinflammatory response that follows TBI can persist for months following injury[58].

TBI often has long term consequences including a potential increased risk of neurodegenerative disorders[1,59–63]. Meta analysis from 32 independent epidemiological studies, totalling “2,013,197 individuals, 13,866 dementia events and 8,166 AD events”, showed TBI increased the risk of any form of dementia by 1.6 times, with individuals showing a 1.5 times higher risk for AD specifically[64]. Many of the proteins associated with neurodegeneration have been shown to accumulate in the brain following TBI, including amyloid beta[65,66], tau[66] and α-synuclein[67]. Chronic traumatic encephalopathy (CTE), a neurodegenerative disorder characterised by the accumulation of hyperphosphorylated tau, has specifically been linked to consistent and repeated brain trauma[68].

Research into the molecular pathways that may drive this connection has suggested that chronic neuroinflammation driven by microglial responses may be responsible for the long term neurodegeneration risk associated with TBI[63,69]. Human brain autopsy samples from patients who have previously experienced a TBI have densely packed, reactive microglia that are not observed within aged matched control

samples[70]. The presence of these reactive microglia also appears to correlate with white matter degeneration, although only observational correlations were provided within this study. While some studies suggest that prolonged activation of microglia has a harmful impact on cognitive function there is also conflicting evidence that microglia may have a neuroprotective effect following TBI[63]. For instance, in a small randomised control study, TBI patients treated with the antibiotic minocycline showed a reduction in microglial activation but an increase in neurodegeneration compared to those patients not given the drug[71]. As well as the conflicting nature of some of the evidence around long-term microglial involvement in TBI, it should also be noted that neither side of the argument provides conclusive proof that microglia functions are driving the potential link between TBI and neurodegeneration.

The epidemiological studies linking TBI to dementia risk can also be difficult to interpret for a variety of reasons including misclassification of neurodegeneration and a lack of official clinical information[63]. It may also be that the link observed between TBI and AD could be driven by hidden factors that increase the risk of both AD and TBI without a causal link between the two. This means further work needs to be carried out on more controlled patient groups in order to fully understand the impact of TBI on dementia risk. It would also be worth building our understanding of how genetic risk factors can impact both TBI outcome and dementia risk. For instance, variants in the APOE gene linked to AD risk have been shown to impact TBI outcomes[72] but the interplay between the two is poorly understood.

1.4.2 Microglia in Multiple Sclerosis

Multiple sclerosis (MS) is a chronic neurological condition that is classified as both a neurodegenerative and autoimmune disorder. The immune system begins to attack the myelin sheath that surrounds neurons in the brain which leads to a multitude of symptoms including muscle weakness and coordination deficits. T-cells, primed to recognise myelin as forgien, are the driving immune cell type behind the development of MS.

While microglia are not associated with the onset of MS, the cells are present in the characteristic brain lesions of MS patients[73,74] and have been shown to be found near

to degenerating neurons in the brain[74]. The presence of the cells within diseased regions and their clear involvement in the immune response in the brain provides some evidence that microglia are involved in disease progression. However, as seen in TBI, different studies report opposing impacts of microglia function: either suggesting they further the progression of MS or that microglia play a neuroprotective role.

Production of reactive oxygen species (ROS) has been implicated in a variety of processes in MS [75] and microglia are often thought of as the major source of ROS within the brain. Microglia within the brain have been shown to express myeloperoxidase (MPO) and generate ROS as part of the myelin phagocytosis process[76]. Expression of MPO also significantly increased in MS patients compared to controls, with the highest level of expression seen in myeloid cells closest to lesion sites. The concept that microglia are the major source of ROS within MS has been further backed-up by more recent experimental data[77] and is thought to be due to Nox2 dependent oxidative burst. Microglia have also been shown to modulate neuronal activity in MS, further adding to described symptoms of the condition. In the Experimental Autoimmune Encephalomyelitis (EAE) mouse model of MS, activated microglia have been shown to release TNFα[78] which can in turn lead to enhanced glutamate function and synaptic degeneration.

On the other hand, a growing body of evidence has linked microglial function to protective disease processes, particularly remyelination[79,80]. *CX3CR1* KO mice, which have altered microglial functions, had a significantly reduced clearance of myelin debris in the EAE model which prevented remyelination[79]. It is also thought that anti-inflammatory microglia can aid the oligodendrocyte differentiation that is required for the remyelination process[80].

1.4.3 Microglial response in other neurological disorders

As the reactive immune cells within the brain, microglia have also been shown to respond to a variety of other neurological disorders, even though they may not play a causal role in the development of the disease. For instance, autism patients have increased microglia cell numbers when compared with healthy controls[81] and have

increased inflammatory profiles within the cerebrospinal fluid, including increased expression of macrophage chemoattractant protein (MCP)-1[82]. Microglia in autistic individuals may also be morphologically distinct. Morgan *et al.*[83] described a reduction in the number and length of distinctive microglial processes within the postmortem tissue from 13 male individuals with autism. Positron emission tomography (PET) scanning has revealed increased levels of microglial activation in autistic brains when compared to healthy controls[84]. Transcriptional profiling of brain tissue from autism patients has highlighted an increased expression of type 1 interferon genes compared to controls[85] and an enrichment of immune module genes within patient samples[86]. However, the genes linked to this immune module showed no enrichment for autism genome-wide association study (GWAS) genes. The lack of enrichment of immune genes within autism GWAS studies implies that the microglial response seen in patients is reactive rather than causal.

Microglia have also been linked to the symptoms associated with neuropathic pain[31,87,88], a chronic and debilitating pain caused by trauma, infection or pathology explicitly linked to peripheral nerve damage. As well as chronic pain symptoms, neuropathic pain also causes tactile allodynia: a disorder when pain hypersensitivity can be caused by what would normally be considered innocuous stimuli. While microglia are not involved in the initial pain stimuli or signalling, they have been shown to react to nerve damage associated with the disorder. Following initial peripheral injury there is marked neuroinflammation, microglial proliferation [89,90] and increased surveillance [91] by microglia. Crosstalk between neurons and microglia, through the CSF-1R signalling pathway, has also been linked to the onset of pain hypersensitivity[92]. Deletion of the *CSF1* gene from sensory neurons, which inhibits production of the signalling molecule, reduced pain hypersensitivity and microglial activation in mice.

1.5 Alzheimer's disease and microglia

Alzheimer's disease (AD) is the most common cause of dementia, a disease that affects around 850,000 people in the UK. Symptoms include progressive memory

loss and a reduction in general cognitive function. AD is also characterised by a general loss of neuronal mass. AD was first described by Dr. Alois Alzheimer in the early 1900s [93,94], where he noted plaques and tangles in patient autopsy samples that are now classically associated with AD pathology. AD is now clinically often split into two distinct categories: familial (early onset) and late onset AD (LOAD). It is thought that early onset AD makes up approximately 5% of all diagnosis' with this branch of the neurodegenerative disorder thought to be highly heritable[95]. Appearance of early onset AD symptoms often occur in patients in their 30s or 40s but are grouped up until the age of 65. Those that appear to sporadically develop symptoms after the age of 65, which is the more common condition, are classified as LOAD patients.

1.5.1 Early hypotheses in Alzheimer's disease research

The first major AD hypothesis focussed on the loss of cholinergic neurons within the brain[96]. Evidence of reduced acetylcholine release and its links with learning and memory further added to the theory[97]. The cholinergic hypothesis was the driver behind major pharmaceutical developments in AD treatments including the cholinesterase inhibitors that are still used in therapy today. However, since their approval as AD therapies, the cholinergic based treatments have appeared to only provide symptomatic relief with little to no effect on the progression of AD[98]. These observations suggest the specific loss of cholinergic neurons may not be driving the progression of the disease.

As understanding of the pathology of AD developed, the amyloid cascade hypothesis became the prevailing pathological theory. The amyloid cascade hypothesis states that it is the formation of the plaque like structures, seen within AD patient brains, that are the molecular drivers of the disease. It is now well accepted that the plaques first described by Alois Alzheimer are made up of aggregated amyloid protein (Aβ), specifically Aβ-42, and neurofibrillary tangles are composed of hyperphosphorylated tau. Hardy and Higgins were the first to coin the "amyloid cascade hypothesis" [99] and put forward the theory that the accumulation of plaques in the brain was the initiating stimulus that led to neuronal loss and the appearance of tau tangles. Since its development, amyloid and its role in the disease has been a major focus of AD research.

The earliest evidence implicating amyloid in AD came from studies of familial AD. Mutations within the amyloid precursor protein (*APP*) gene[100,101] and within the presenilin genes *PSEN1* and *PSEN2*[101–103] cause familial AD. APP, PSEN1 and PSEN2 are all involved in the production of the toxic Aβ-42 protein that forms the major component of plaques. The APP protein can be cleaved in different ways that lead to the production of a variety of forms of amyloid beta. It is thought that mutations associated with familial AD cause a bias towards the cleavage mechanism that generates the toxic Aβ-42. Further support came from early onset of AD in patients with Down's syndrome, who have three copies of the APP gene[104]. While mice do not spontaneously develop AD-like pathology or symptomatology as they age, *APP* and *PSEN* mutant mice have been shown to develop cognitive deficits, amyloid accumulation and synaptic loss[104].

Since the initial description of the amyloid cascade hypothesis, large bodies of research using a variety of molecular tools have been used to demonstrate that various forms of Aβ can initiate symptoms of AD[104–106]. For instance: in rat hippocampal cultures the addition of aggregated Aβ is neurotoxic[107]**,** APP transgenic mice have increased levels of Aβ oligomers and the same miceshow significant cognitive impairment compared to controls[108]. In mouse models of AD disrupting the amyloid pathway can result in a reversal of many of the cognitive phenotypes seen in the mice[109,110].

The growing evidence from *in-vitro* and *in-vivo* studies led to a push for drugs targeting the amyloid pathway. However, the amyloid cascade hypothesis is not without controversy [104,111,112]. One of the most significant problems with the theory that amyloid is the driver behind AD pathology is the repeated failure of anti-amyloid therapies in clinical trials[113]. These therapies fall into two broad categories: direct reduction of Aβ through antibody-style therapies and targeting of enzymes involved in the production of amyloid, such as BACE and γ-secretase. Many of the drugs targeting the enzymatic pathways have failed in clinical trials, either due to lack of efficacy[114] or significant off-target effects[115,116].

Despite the initial clinical safety failings of immunotherapies targeting Aβ[117], multiple therapies reached phase II and III trials[104]. However, the majority of these compounds have also dropped out of trials due to the failure to meet clinical endpoints[113]. In 2014, data was published from phase III trials of the anti-Aβ monoclonal antibody Bapineuzumab in which patients on the drug showed no significant improvement in AD-linked cognitive function compared to the placebo group[118]. The failure of Bapineuzumab in phase III trials came despite evidence from earlier phase II studies that long term treatment with the drug significantly reduced cortical amyloid fibrillar load[119].

The fact that immunotherapies targeting the amyloid pathway appear not to halt disease development despite reductions in amyloid load, has led to suggestions that targeting amyloid is the wrong strategy since it is not driving AD progression[113,120]. It is worth noting, however, that in late 2019 pharmaceutical company Biogen announced that they were seeking FDA approval for their anti-Aβ antibody despite earlier failure of the drug in trials[121]. The repeated failure of AD modifying drugs in clinical trials leads to questions not just about the validity of the targets but also practical factors about how trials are carried out[120] including whether patients are targeted for treatment too late in disease progression. There are also questions around the sensitivity of the major cognitive test used in AD clinical trials, the Alzheimer's Disease Assessment Scale–Cognitive Subscale (ADAS-cogs), particularly in the early mild stages of disease[122].

1.5.2 Alzheimer's disease genetics and the neuroinflammation hypothesis

Although the amyloid cascade hypothesis has driven a large part of AD research, it is important to remember that the theory was founded on the genetics of early onset, familial, AD. The genetics behind LOAD is more complex and heterogeneous, not driven by single mutations in disease linked genes but by large numbers of variants of individually small effect sizes.

One of the first major genetic risk factors that was identified in LOAD is the *APOE* gene, a protein involved in cholesterol transport[123,124]. Specifically it has been demonstrated that the ε4 allele significantly increases AD risk, while the ε2 allele

confers a protective effect compared to the other alleles[125]. Early studies of the genetic risk factors for LOAD were carried out in relatively small patient numbers. This only allowed for the identification of single nucleotide polymorphisms (SNPs) which conferred relatively large increases in risk, such as *APOE*, or those within small targeted gene sets identified before analysis, such as *SORL1*[126]. However, genome-wide association studies (GWAS) have generated large scale datasets from case/control comparisons that can detect small effect size genetic links to complex disorders including AD[127].

While activation of the immune system, particularly microglia, was known to occur in AD as part of normal pathology[128,129], for many years this was thought to be a downstream effect of the disease. The results of AD GWAS provided the first indication that the innate immune system may have a causal role in the development of AD. Identification of SNPs near genes such as *CD33*, *CR1* and *MS4A6A*, which are classically considered immune related, suggests some role for the immune system within the disease. The identification of rare missense variants in genes, such as *TREM2*, *ABI3* and *PLCG2*, which are highly expressed in immune cells[130] has provided further evidence for the neuroinflammation theory. Table 1.1 lists the risk alleles identified in AD GWAS studies and the nearest gene to each SNP.

Lead SNP	Nearest gene	Publications
rs3851179	*PICALM*	131–134
rs10792832		135–137
rs11136000	*CLU*	131–133,138
rs9331896		134–136
rs4236673		137
rs3818361	*CR1*	132,139
rs6656401		135–138
rs4844610		134
rs744373	*BIN1*	133,139
rs6733839		134–136
rs4663105		137
rs3764650	*ABCA7*	139

rs4147929		[135,136]
rs3752246		[134]
rs111278892		[137]
rs610932	*MS4A6A*	[139]
rs983392		[135,136]
rs2081545		[137]
rs10948363	*CD2AP*	[135,136]
rs9473117		[134]
rs9381563		[137]
rs11771145	*EPHA1*	[135,136]
rs10808026		[134]
rs7810606		[137]
rs3865444	*CD33*	[135–137]
rs28834970	*PTK2B*	[135,136]
rs73223431		[134]
rs11218343	*SORL1*	[134–137]
rs10498633	*SLC24A4*	[135,136]
rs12881735		[134]
rs12590654		[137]
rs8093731	*DSG2/SUZ12P1*	[135,137]
rs35349669	*INPP5D*	[135,136]
rs10933431		[134,137]
rs1476679	*ZCWPW1*	[135,136]
rs1859788		[137]
rs17125924	*FERMT2*	[134–136]
rs7274581	*CASS4*	[135,136]
rs6024870		[134]
rs6014724		[137]
rs593742	*ADAM10*	[136]
rs442495		[137]
rs889555	*BCKDK/KAT8*	[136]
rs59735493		[137]
rs138190086	*ACE*	[134,136]
rs12444183	*PLCG2**	[136]
rs75932628	*TREM2**	[134]

rs187370608		[137]
rs7920721	*ECHDC3*	[134]
rs11257238		[137]
rs28394864	*ABI3**	[137]
rs179943	*ATXN1*	[140]
rs3826656	*NT_011109.848*	[140]
rs2049161	*BC040718*	[140]
rs597668	*EXOC3L2*	[133]
rs670139	*MS4A4E*	[139]
rs190982	*MEF2C*	[135]
rs2718058	*NME8*	[135]
rs10838725	*CELF1*	[135]
rs9381040	*TREML2*	[136]
rs59685680	*SPPL2A*	[136]
rs4985556	*IL-34*	[136]
rs3740688	*SPI1*	[134]
rs7933202	*MS4A2*	[134]
rs4575098	*ADAMTS4*	[137]
rs184384746	*HSEX1*	[137]
rs6448453	*CLNK*	[137]
rs114360492	*CNTNAP2*	[137]
rs117618017	*APH1B*	[137]
rs113260531	*SCIMP*	[137]
rs2632516	*BZRAP1-AS1*	[137]
rs76726049	*ALPK2*	[137]
rs76320948	*AC074212.3*	[137]

Table 1.1 Summary of reported AD GWAS hits

Lead SNPs and nearest genes identified in AD GWAS studies. Certain loci have differing lead SNPs identified by studies but are grouped by nearest gene. Loci with a * next to the gene name have previously been identified in rare variant studies.

The results of GWAS studies displayed here provide summaries of each locus, highlighting only the most associated SNP and the nearest gene to that SNP for each region. Linkage-disequilibrium (LD) within the human genome is a terminology that describes certain SNPs within a region that are found to be more associated with

each other than would be expected if they were inherited randomly. This means there are often multiple SNPs within a region in strong association with the “lead” SNP identified in a GWAS. It is, therefore, not possible to tell from standard GWAS analysis which of these SNPs is causal. Additionally, because disease associated variants are noncoding, there are many genes within a specific window of the associated SNPs that could be impacted by the variant. This means that it is also not possible to tell exactly which gene, and downstream signalling pathways, may be linked to disease risk.

To address these problems, methods to combine GWAS data with functional data, including transcriptomics (expression quantitative trait loci (eQTL) maps) and open chromatin assays (chromatin accessibility quantitative trait loci (caQTL) maps). It is then possible to run co-localisation analysis to identify variants affecting both disease risk and a functional output have been developed. Computation tools also provide methods to extend traditional GWAS analysis. For instance, GoShifter[141] prioritises functional annotations to identify causal variants by finding SNP enrichments in annotated regions.

In the case of AD, these combination approaches have further linked the immune system to disease risk. For instance, when eQTL maps of monocytes and T cells were colocalized with GWAS summary statistics from a variety of complex traits, significant co-localisations with AD GWAS SNPs were only identified within the monocyte eQTL map[142]. While this implied that the myeloid cell lineage of the immune system may be driving the neuroinflammatory component of AD, it did not fully rule out a role for neurons themselves. Integrative analysis of published GWAS summary statistics and whole-brain single cell RNA-sequencing data shows a significant enrichment of AD GWAS signal within the specific gene expression pattern of microglial cells, while no enrichment was seen in neurons[143]. AD risk SNPs are also significantly enriched in regions of open chromatin in myeloid cells, including microglia, but not in whole brain chromatin accessibility data[144]. Although AD genetics studies have now identified multiple risk loci these have not yet provided direct information on the biological role of microglia in neurodegeneration.

1.5.3 The role of microglia in Alzheimer's disease

Genetic studies have spurred a resurgence of research into how microglial function changes during AD. When Alois Alzheimer first described the brain pathology of AD, in addition to identifying amyloid plaques and tau tangles, he also observed alterations in the glia surrounding these abnormal proteins, including the development of "fibers" and "adipose saccules"[93]. Since this initial description, there has been a growing body of research that focuses on microglial involvement in AD. This has provided evidence that often falls into one of two categories: that promoting microglial activity will be beneficial in AD or that a reduction in activity will slow AD progression. However, these two ideas may not be mutually exclusive in that certain processes may be both beneficial or harmful depending on the context.

Microglial phagocytosis is a good example of the above phenomenon. Initially, research focussed on microglial phagocytosis of amyloid plaques within the brain[129,145,146], in part due to the observed physical association of microglia with the plaques. It has been suggested that microglial recruitment to plaque sites promotes phagocytosis and lowers plaque burden[147]. However, as the disease progresses the phagocytic capability of microglia reduces[148] and in fact the cytokines produced by the process are part of a negative feedback loop that reduces phagocytosis[147]. The evidence from these mouse studies implies that promoting microglial phagocytosis could be a viable therapeutic target as it reduced amyloid load. However, selective reduction in microglial populations in an AD mouse model may reduce neuronal loss without impacting amyloid load[149] which suggests microglial phagocytosis of amyloid is not necessarily required for the reversal of AD symptoms. In fact, microglial phagocytosis, via a complement dependent mechanism, has since been linked to excessive engulfment of healthy synapses[150]. This means that increasing microglial phagocytic capabilities may in turn lead to further neuronal loss.

Outside of phagocytosis, microglia have been linked to a variety of other molecular processes in AD. For instance CSF-1R inhibition in the 5XFAD mouse model of AD has been shown to significantly reduce the seeding of plaques within the brain[151,152], although Aβ accumulation still appears in cortical blood vessels. Other work suggests that microglia may form a barrier around developing plaques which reduces further

accumulation of Aβ[146]. In tauopathy mouse models, microglia aid the propagation of tau across the brain via the secretion of previously phagocytosed tau in exosomes[153].

Further insights into microglial functions in AD have come from studying mutations identified by GWAS. For example, multiple studies have functionally characterised mutations in *TREM2*[154,155]. Triggering receptor expressed on myeloid cells 2 (TREM2) is a receptor that signals through a TYROBP/DAP12 dependent mechanism to activate a variety of signalling pathways and downstream functions, such as phagocytosis and chemotaxis[156]. A variety of approaches have shown that disease-associated missense mutations in *TREM2* can alter microglial phagocytosis, survival and proliferation[156]. The soluble form of TREM2, produced following cleavage of the receptor, has also been implicated in AD[157–159]. There is evidence that TREM2 may function in conjunction with other GWAS risk genes during AD including *APOE*[160,161], *CD33*[162] and *MS4A*[163].

Alternative experimental approaches have examined how microglial functions change in AD patients compared to age matched healthy controls, particularly at the level of gene expression. In mice, two studies have identified microglial populations that only appear in diseased states[164,165] and identify a loss of homeostatic gene expression (*P2RY12*, *CX3CR1* and *TMEM119*) alongside an increase in inflammatory markers such as *AXL*, *CLEC7A* and *CST7*. Additionally, activation of TREM2 signalling pathways were required for the formation of this disease associated subtype of microglia cells in mice. In human samples, single cell analysis of AD post-mortem brain samples also identified a disease specific population of microglial cells[166]. Like the populations identified in mice, these cells had increased expression of genes like *SPP1* and *APOE*. The disease specific microglia also showed an increased expression of HLA and complement linked genes, compared to non-disease linked microglia.

In summary, it is clear that microglia play a significant role in how our brains function in health and disease but exactly how microglial processes change in disease and precisely how to target the same pathways in treatments remains unclear. Much of this complexity often arises because microglia seem to play both detrimental and

beneficial roles in many diseases depending on the stage, activation pattern or model system being studied.

1.6 Studying human microglia

While significant advances have been made in microglial research, many of the studies that have been used to understand microglia function in health and disease have been carried out in mice. Mouse models are an invaluable tool, enabling large scale studies, manipulation of the cells and providing a way to study microglia throughout the lifespan of an organism. However, studies in mice are not without limitations and controversies[167–169]. There are significant differences in the fundamental functions of microglia in mice and humans, including differences in marker expression, such as IFNγ and TLR4, and differences in response to pharmacological compounds. In mouse models of AD, microglia are often described as taking on an activated phenotype while in human autopsy samples the cells appear to degenerate with age, often referred to as dystrophic or senescent[170]. This can lead to opposing theories about the role microglia play in disease.

However, primary human microglia are extremely difficult to source and come with experimental caveats. Many commercially available human microglia sources are fetal samples which may behave differently to fully developed microglia. Additionally, commercially available cells are often cultured which can impact microglial expression[171]. Protocols for accessing human adult microglia cells from both post-mortem and surgical tissues have been refined and appear to yield relatively pure samples[172–174]. Although isolated human microglia may have high purity, there are multiple experimental factors to consider when using these cells. Even small periods of culturing can alter the profile of human microglia[171,175] and little is known about how the isolation protocols (dissociation and cell marker expression based sorting) may impact microglial profiles. Small scale microarray analysis of sorted murine mammary glands has suggested that fluorescence activated cell sorting (FACS) has minimal impact on gene expression[176]. However, full comparisons have

not been carried out to understand how FACS sorting may impact immune cell expression, particularly microglia.

While it is possible to isolate fresh primary adult human microglia from neurosurgical patients, in order to study microglia from healthy individuals, samples must be acquired from post-mortem tissue. As microglial phenotypes have been shown to be heavily dependent on the active neuronal environment[171], it is therefore difficult to know how much post-mortem delay impacts microglia. A study comparing isolated microglia from brains with differing lengths of post-mortem delay demonstrated that disease state had a greater impact on microglia than the time between death and collection[175]. However, it is difficult to directly compare fresh microglia to post-mortem samples while controlling for confounding factors. Therefore, it is impossible to definitively know the impact of post-mortem collection on microglial phenotype.

1.6.1 Transcriptomic studies in primary human microglia

RNA-sequencing technology enables the study of the whole transcriptome of cells and whole tissues. Statistical analysis of the resulting data can be used to compare the transcriptional profiles of samples across a variety of conditions. As isolation protocols for human primary microglia have improved RNA-sequencing has become widely used to understand differing aspects of microglia. This includes comparisons between human and mouse samples[171], identifying microglia-specific marker genes[177,178], comparison of transcriptomes across ages[179], highlighting region and disease specific changes in gene expression[180] and understanding the role environment plays in microglial gene expression[171].

While RNA-sequencing at a bulk level has provided tools to study large scale gene expression and generated vast amounts of data, the ability to use the technology at the single-cell resolution has provided a tool to study gene expression at a much finer resolution[181,182]. Single-cell RNA-sequencing (scRNA-seq) allows identification of individual populations of cells *in silico,* obviating the need for prior knowledge of cell markers, and enabling comparisons of tissue composition between experimental groups.

scRNA-seq has allowed researchers to take whole brain tissue and identify multiple cell types, such as neurons and microglia[166,183,184]. Whole brain single cell analysis has been used to investigate changes that occur to different cell types in the brain during development[183] and disease[166,184]. Being able to identify microglia from whole brain samples also removes the cell sorting step required for bulk RNA-sequencing, which in turn reduces the chances of experimental processes impacting microglial gene expression. However, within whole brain single cell analysis the fraction of microglia is relatively low (3% reported by Mathys *et al.*[166]) and smaller numbers of cells per subgroup makes statistical comparisons more difficult. Therefore, it is also possible to use single-cell sequencing on sorted primary human microglia[185,186], in order to better capture subtle microglial population changes. This has been used to further our understanding of microglial populations across ages[185] and disease[185,186].

An extended review of how transcriptional analysis of primary microglia has impacted our understanding of the cell type can be found in section 2.1. While current published datasets have provided an insight into microglial transcriptomes, many are still based on relatively small patient numbers. This is largely because access to primary human microglial samples is still difficult. Growing brain bank collections have allowed access to larger numbers of post-mortem samples but these studies are still limited by patient number (with the largest reported at 48 collections[166]) and often cover only specific disease states. Fresh human microglia are even more difficult to access, coming from either fetal samples or neurosurgical patients.

1.6.2 Modelling human microglia

While studying primary human microglia is important for understanding the cells in health and disease, there are clear limitations with these studies particularly around scale and the ability to experimentally manipulate the cells. Therefore, a clear challenge has been to develop ways to model human cells in the lab. Induced pluripotent stem cells (iPSCs) are proliferating cells that have been reverted back to a stem cell like state from adult cells and they have the potential to differentiate into any cell[187–190]. This means iPSC based cell model systems provide researchers with a useful tool for studying human disease in a dish[191]: they are able to be used at scale, can be manipulated experimentally and allow for repeated sampling. Large scale

banks of iPSC lines, such as the HipSci consortium, mean that researchers can also run iPSC based experiments using large numbers of both healthy and diseased cell lines.

As iPSC cells can technically be differentiated into any cell in the body, methods have been developed to differentiate these cells along a myeloid lineage. Initially these studies focussed on the development of macrophage models and their utilisation for studying immune response[192–195]. Many of these iPSC-derived macrophage differentiation protocols make use of the induction of embryoid bodies (EB) from stem cells. These EB structures are made up of cells from all three germ layers[196] that can then further differentiate into more specialised cells.

However, in more recent years there has also been a focus on pushing the myeloid cells closer towards the specialised microglia-like phenotype. These protocols range from simple monoculture based systems[197–201], similar to those used to generate macrophage-like cells, to more complex co-cultured[198,202] and organoid systems[200,203–206]. These more complex model systems build on the idea that much of the unique microglial transcriptional signature comes from the environmental stimulation they receive from neurons and other parts of the CNS[171].

A major factor to consider when using *in-vitro* models for human cells is understanding how accurately the cell culture systems capture the primary cell type. Often this comparison is limited to marker gene expression and functional capabilities. For a detailed analysis of how the iPSC models described above have been compared to primary cells see Chapters 3 and 4. For microglia particularly, comparison is complex, as the primary cells are difficult to access and therefore transcriptional comparisons are often made across studies. This can often lead to confounding batch effects, especially when running small scale comparisons. Systematic comparisons of model systems to the primary microglia can be used to highlight potential signalling pathways that are not switched on *in-vitro* and could be manipulated to move cells closer towards the primary cell type.

1.7 Thesis overview

The overarching theme of the following thesis builds on section 1.6 and the difficulties around studying human microglia. I aim to answer three major questions throughout the thesis: **1.** How does microglial composition and gene expression profile change across a population? **2.** How accurately do current simple *in-vitro* model systems of human microglia capture the profile of primary human cells? **3.** Does culturing stem cell derived microglia with neurons move the model systems closer to the primary phenotype?

The analysis in the second chapter of my thesis forms part of a large-scale project in collaboration with Dr Adam Young and Dr Natsuhiko Kumasaka studying the genetic architecture of human primary microglia. As part of the project we collected and processed the largest number of fresh, primary human microglia samples to date from a wide variety of clinical phenotypes. In this chapter I used single cell RNA-sequencing to identify different subpopulations of primary microglia and identified how the likelihood of finding cells within these populations is influenced by clinical phenotypes. I then used bulk and single cell RNA sequencing data from the same patient population to further understand how clinical phenotypes such as age, pathology and sex influcenced microglial transcriptomes.

In the third chapter of my thesis, I focus on the transcriptional profiles of *in-vitro* models of microglia and how closely they match the transcriptional profile of the primary human cell type. I collected publicly available data and combined it with available in-house datasets to generate a large scale analysis project to compare primary human microglia with monocyte-derived macrophages, cancer-cell lines, iPSC-derived macrophages and iPSC-derived microglia. For all the data, I used raw sequencing files that were all processed through the same pipeline and I ensured that I collected data from multiple studies for each cell type. Both of these decisions were made to reduce the batch effects that can occur when comparing sequencing data across different studies[207–209]. I used the processed data to understand how the

different *in-vitro* systems capture the gene expression of the primary cells and which signalling pathways may not be switched on in these *in-vitro* systems.

In the final results chapter of my thesis, I will focus on more complex stem cell based model systems, including co-culture and orgainoid based models. This forms part of a collaboration with Dr Phil Brownjohn and Dr Moritz Haneklaus, from the Livesey Lab, working with their published microglia differentiation protocols[200]. I initially used bulk RNA-sequencing to add the complex model systems to the large dataset generated in Chapter 3 in order to understand how the more complex model systems compared to the monoculture systems described in Chapter 2. I then used single cell sequencing, and particularly single cell trajectory analysis, to understand how microglial cells from each of the model systems fit on a developmental pathway that ultimately ends with the primary cell type.