

Title Page

Title:

Variation on a theme: Mapping microglial heterogeneity

Authors:

Sam J Washer^{1,2}, Andrew R Bassett², Sally A Cowley¹

¹ James Martin Stem Cell Facility, Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, OX1 3RE

² Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, CB10 1SA

Corresponding Author

Dr Sally A Cowley

James Martin Stem Cell Facility

Sir William Dunn School of Pathology

University of Oxford

South Parks Road

Oxford

OX1 3RE

Email: sally.cowley@path.ox.ac.uk

Key Words:

Microglia, single cell RNA sequencing, transcriptomics, genetics, stem cell, modelling

Abstract

Young et al examine the complexity of primary human microglia, identifying previously unknown cell states. Using eQTL mapping techniques, they identify 129 genes whose expression in microglia is linked to disease, and show that iPS cell models can be used for functional validation of common genetic mutations in microglia-associated diseases.

Main

Microglia are the major resident macrophages of the brain, and are responsible for a variety of normal homeostatic functions, including synaptic pruning for neuronal plasticity, and clearance of dead cells and misfolded proteins (reviewed in [1]). As a result they are implicated in a number of neurodegenerative disorders, particularly Alzheimer's disease (AD) [2], Parkinson's disease (PD) [3] and amyotrophic lateral sclerosis (ALS) [4]. However, profiling microglia from patients at post-mortem alone captures the final endpoint of the disease, and such studies are further confounded as brain samples are inevitably frozen with a range of post mortem intervals, which impacts the microglial transcriptome [5]. Thus the vast majority of our knowledge about microglial heterogeneity/characterisation comes from unreliable tissue from small, underpowered studies (Table 1).

Young et al. set out to address these problems by performing the largest transcriptome study to date by using fresh microglial samples from a large number of patients undergoing neurosurgery (141) [6]. Their cohort contained a near even split of genders with a range of ages, brain regions, and conditions (hydrocephalus, tumours, haemorrhage, and trauma). They isolated microglia from whole brain using CD11b-

coated magnetic beads or fluorescence-activated cell sorting and undertook bulk and single cell RNA-seq, and used blood for genotyping (summarised in Figure 1).

From the single cell RNA-seq they identified four distinct populations of microglia, which they characterise as naïve, sub-acute activation, and two states of acute activation. The naïve population have an upregulation of known, well characterised microglia marker genes (*P2RY12/CX3CR1/TREM2*), while the other three populations show high expression levels of immune response and activation genes (*IL1b/CD83/CCL3*). The microglia from haemorrhage and trauma patients (which have not been previously characterised) display genes involved with activation of acute immune response (*NF-KB/STAT3/RUNX1/MHC-I*). The authors confirm this by comparing their subtype marker genes to markers from AD-associated microglia, finding that they closely resemble the naïve and sub-acute populations. Interestingly, biological age, one of the biggest risk factors for AD, has a large effect on gene expression in their dataset, with microglia from older patients exerting increased inflammation profiles, and decreased motility and proliferation. However most of the variation between the microglial samples is driven by patient identity, suggesting genetic causality or sampling batch variation.

To examine the effect of genetics on transcriptional response, they identify expression quantitative trait loci (eQTL). eQTLs are a naturally occurring haplotype, associated with a change in gene expression, likely through one or more of the SNP in that haplotype promoting/inhibiting transcription factor binding or altering chromatin confirmation [7]. The authors exclude samples with non-European ancestry (to minimise differences in linkage disequilibrium), resulting in 93 individuals for eQTL mapping, within which they identified 585 eQTLs where a haplotype is predictive of a gene expression change. As eQTLs have tissue specific effects, they looked at shared

eQTLs across different cell subtypes commonly used to study microglia *in vitro* – blood-derived monocytes, and induced Pluripotent Stem Cell (iPSC) derived macrophages (iPSDmac). This identifies 108 microglial specific eQTLs, 449 across all cells types, 192 shared between microglia and iPSDmac, and 106 between microglia and monocytes. Colocalization of the eQTLs with 146 previous genome wide association studies identified 129 genes and shows that the strongest associations are with AD/PD/irritable bowel disease, diseases with a known microglial association.

Of the eQTLs identified in this study, 11 have been previously associated with AD. Of these, 3 are present in iPSDMacs (*BIN1/EPHA1/PTK2B*), a cell type much easier to derive than primary microglia, offering a potential alternative cell model for studying these mutations. But is the *direction* of effect shared between the two models? The authors discovered that this is not always the case - for example the *PTK2B* eQTL has the opposite direction of effect. However this is not the case with *BIN1*, where the direction of effect is the same in both microglia and iPSDMacs. With *BIN1* as an example, they undertake in-depth analysis of the chromatin structure at this *locus* using 5 primary microglia lines, 89 iPSDMac cell lines and ATAC-Seq (which identifies regions of open chromatin), to establish whether the eQTL is affecting gene expression through chromatin changes. Excitingly, they show that the eQTL sits within a region of open chromatin in both primary microglia and iPSDMac, indicating that this eQTL can be studied using this cell line. Furthermore, they uncover that the eQTL is a predicted binding site of the transcription factor MEF2C. They conclude that in primary microglia and the iPSDMac model of microglia the same genetic signal enhances AD risk, increases chromatin accessibility by enhancing binding of MEF2C, and results in increased *BIN1* expression. As further confirmation, co-expression of *BIN1* and *MEF2C* is found only in primary microglia and iPSDMac. This is important

as *BIN1* is one of the most common genetic variants in AD, and this shows that it can be studied using a relatively simple cell model, iPSDMac (as used in [8]).

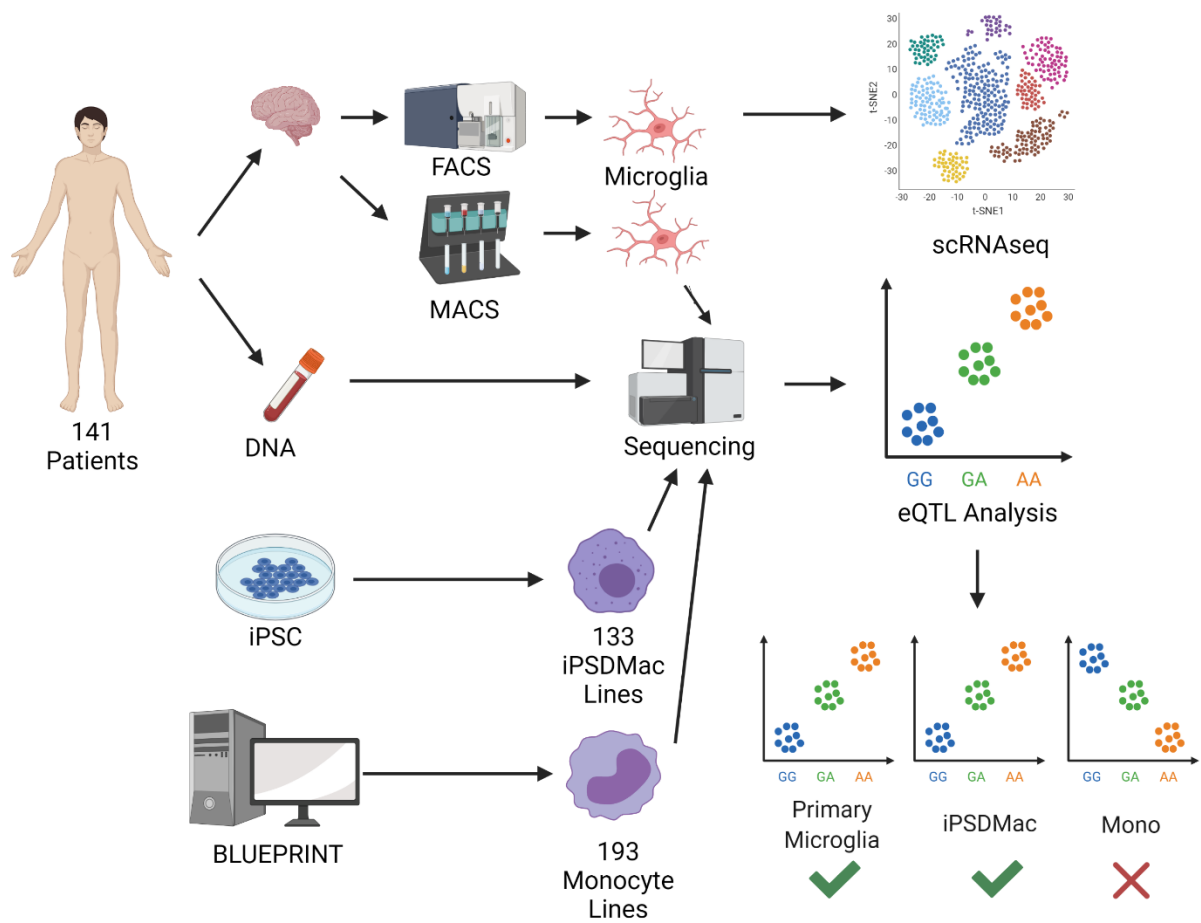
Another exciting aspect of this study is the identification of previously unknown early activation states of human microglia, which should help us understand the earliest stages of some of the more complex neurological diseases. This further highlights that microglial activation is more complicated than a simple generic proinflammatory response to any stimuli, but that they can respond specifically according to the type of stimulus. We can now start to examine how the presence of misfolded proteins involved in neurodegenerative diseases, particularly Amyloid-beta, Microtubule-associated protein tau, Alpha-synuclein, overlaps with these early (albeit acute) activation phenotypes, and examine disease progression at the earliest stages. The other interesting outcome of the study is the use of human genetics to identify 129 genes, where shared genetic variants cause a change in gene expression in microglia as well as increase risk of disease. They are also able to assess the best cell models to elucidate mechanisms. Similar eQTL mapping methodology can also be used for other diseases, especially PD, ALS and ALS, to identify target genes having their effects in microglia and assessing the microglial states that are important in driving disease.

Figure 1: Summary of how Young et al. integrate genomic data to uncover previously unknown microglial activation states through single cell RNA sequencing (scRNAseq), identify microglial specific expression quantitative trait loci (eQTLs), and how cell models of microglia can be used to model genetic perturbations in microglia-associated diseases.

Table 1: Summary of previous single cell RNA sequencing experiments performed on primary microglia.

References

1. Prinz M, Jung S, Priller J. Microglia Biology: One Century of Evolving Concepts. *Cell* [Internet]. 2019;179(2):292–311. Available from: <https://doi.org/10.1016/j.cell.2019.08.053>
2. Wightman DP, Jansen IE, Savage JE, Shadrin AA. Largest GWAS (N = 1 , 126 , 563) of Alzheimer ' s Disease Implicates Microglia and Immune Cells Author List and affiliations. *medRxiv*. 2020;1–24.
3. Bryois J, Skene NG, Hansen TF, Kogelman LJA, Watson HJ, Brueggeman L, et al. Genetic Identification of Cell Types Underlying Brain Complex Traits Yields Novel Insights Into the Etiology of Parkinson's Disease. *Nat Genet*. 2020;in press.
4. Clarke BE, Patani R. The microglial component of amyotrophic lateral sclerosis. *Brain* [Internet]. 2020 Dec 1;143(12):3526–39. Available from: <https://academic.oup.com/brain/article/143/12/3526/6016135>
5. Dachet F, Brown JB, Valyi-Nagy T, Narayan KD, Serafini A, Boley N, et al. Selective time-dependent changes in activity and cell-specific gene expression in human postmortem brain. *Sci Rep* [Internet]. 2021;11(1):1–11. Available from: <https://doi.org/10.1038/s41598-021-85801-6>
6. Young AMH, Kumasaka N, Calvert F, Hammond TR, Knights A, Panousis N, et al. A map of transcriptional heterogeneity and regulatory variation in human microglia. *Nat Genet* [Internet]. 2021;53(6):861–8. Available from: <http://dx.doi.org/10.1038/s41588-021-00875-2>
7. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. *Nature* [Internet]. 2017 Oct 12;550(7675):204–13. Available from: <http://www.nature.com/articles/nature24277>
8. Alasoo K, Rodrigues J, Mukhopadhyay S, Knights AJ, Mann AL, Kundu K, et al. Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nat Genet* [Internet]. 2018;50(3):424–31. Available from: <http://dx.doi.org/10.1038/s41588-018-0046-7>



Author	Year	PMID	Species	Age	Disease	Source	Sample	Sorted	Size	Nuclear or Cell	Method
Young et al	2021	34083789	Human	Adult	Healthy, hydrocephalus, tumour, haemorrhage, trauma	Surgery	Cerebellum, Frontal, Occipital, Parietal, Temporal	CD11b+	141	Cell	Smart-seq2
Morabito et al	2021	34239132	Human	Adult	Healthy, Alzheimer's disease (AD)	Post-mortem	Whole brain	No	18	Nuclear	10x Genomics
Marinaro et al	2021	bioRxiv	Human	Adult	Healthy, AD	Post-mortem	Whole brain	NeuN-	16	Nuclear	10x Genomics
Olah et al	2020	33257666	Human	Adult	Healthy, epilepsy	Post-mortem/Surgery	Dorsolateral prefrontal cortex, temporal cortex	CD11b+, CD45+	17	Cell	10x Genomics
Masuda et al	2019	30760929	Mouse	E16.5-16W	Healthy, FNX treated	Post-mortem	Forebrain, midbrain, cerebellum, spinal chord, cortex, corpus callosum, hippocampus, facial nucleus	Cd45+, Cd11b+, Cd206-	60	Cell	Smart-seq2
			Human	Adult	Healthy, multiple sclerosis	Surgery	Temporal, frontal, parietal	CD45+	10	Cell	Cel-Seq2
Hammond et al	2019	30471926	Mouse	E14.5-P540	Healthy	Post-mortem	Whole brain	Cd45 low, Cd11b high, Cx3cr1 high	50	Cell	10x Genomics
Mathys et al	2017	29020624	Mouse	Adult	Healthy, AD Like (CK-p25 induced neurodegeneration)	Post-mortem	Hippocampus	Cd11b+, Cd45+	25	Cell	Smart-seq2
Keren-Shaul et al	2017	28602351	Mouse	7W-8M	Healthy, AD Like (5XFAD), AD Like (TREM KO), ALS Like (mSOD1 G93A)	Post-mortem	Whole brain, cortex, cerebellum, spinal chord	Cd45+, Cd11c, Cd11b+	30	Cell	MARS-seq