Engrafted parenchymal brain macrophages differ from host microglia in transcriptome, epigenome and response to challenge

Anat Shemer[1,#], Jonathan Grozovski[1,#], Tuan Leng Tay[2,3,#], Jenhan Tao[4], Alon Volaski[1], Patrick Süß[3], Alberto Ardura-Fabregat[3], Mor Gross-Vered[1], Jung-Seok Kim[1], Eyal David[1], Louise Chappell-Maor[1], Lars Thielecke[5], Christopher K. Glass[4], Kerstin Cornils[6], Marco Prinz[3,7] and Steffen Jung[1 *]

[1] Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel

[2] Cluster of Excellence BrainLinks-BrainTools, [3] Institute of Neuropathology, Medical Faculty, University of Freiburg, 79110 Freiburg, Germany

[4] Department of Cellular and Molecular Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0651, USA

[5] Institute for Medical Informatics and Biometry, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[6] University Medical Center Hamburg-Eppendorf, Department of Pediatric Hematology and Oncology, Division of Pediatric Stem Cell Transplantation and Immunology and Research Institute, Children's Cancer Center Hamburg, 20246 Hamburg, Germany

[7] BIOSS Centre for Biological Signalling Studies, University of Freiburg, 79106 Freiburg, Germany

# These authors contributed equally to the work

[*] Corresponding author e-mail: s.jung@weizmann.ac.il (S.J.)

Abstract

Microglia are yolk sac-derived macrophages residing in the parenchyma of brain and spinal cord, where they interact with neurons and other glial cells by constantly probing their surroundings with dynamic extensions. Following different conditioning paradigms and bone marrow (BM) / hematopoietic stem cell (HSC) transplantation, graft-derived cells seed the brain and persistently contribute to the parenchymal brain macrophage compartment. Here we establish that these cells acquire over time microglia characteristics, including ramified morphology, longevity, radio-resistance and clonal expansion. However, even following prolonged CNS residence, transcriptomes and epigenomes of engrafted HSC-derived macrophages remain distinct from yolk sac-derived host microglia. Furthermore, BM graft-derived cells display discrete responses to peripheral endotoxin challenge, as compared to host microglia. Also in human HSC transplant recipients, engrafted cells remain distinct from host microglia, extending our finding to clinical settings. Collectively, our data emphasize the molecular and functional heterogeneity of parenchymal brain macrophages and highlight potential clinical implications for patients treated by HSC gene therapy.

Introduction

Macrophages were shown in the mouse to arise from three distinct developmental pathways that differentially contribute to the respective tissue compartments in the embryo and adult. Like other embryonic tissue macrophages, microglia first develop from primitive macrophage progenitors that originate in the mouse around E7.25 in the yolk sac (YS), are thought to be independent of the transcription factor (TF) Myb and infiltrate the brain without monocytic intermediate [1-3]. YS macrophage-derived microglia persist throughout adulthood. Most other tissue macrophages are however shortly after replaced by fetal monocytes that derive from myb-dependent multipotent erythro-myeloid progenitors (EMP) that also arise in the YS, but are currently thought to be consumed before birth. Starting from E10.5, definitive hematopoiesis commences with the generation of hematopoietic stem cells (HSC) in the aorto-gonado-mesonephros (AGM) region. HSC first locate to the fetal liver but eventually seed the bone marrow (BM) to maintain adult lymphoid and myeloid hematopoiesis. Most EMP-derived tissue macrophage compartments persevere throughout adulthood without significant input from HSC-derived cells. In barrier tissues, such as the gut and skin, but also other selected organs, such as the heart, HSC-derived cells can however progressively replace embryonic macrophages involving a blood monocyte intermediate [4].

Differential contributions of the three developmental pathways to specific tissue macrophage compartments seem determined by the availability of limited niches at the time of precursor appearance [5]. In support of this notion, following experimentally induced niche liberation by genetic deficiencies, such as a Csf1r mutation, irradiation or macrophage ablation, tissue macrophage compartments can be seeded by progenitors other than the original ones [6-9].

Tissue macrophages display distinct transcriptomes and epigenomes [10,11], that are gradually acquired during their development [12,13]. Establishment of molecular macrophage identities depends on the exposure to tissue specific environmental factors [4,14]. Accordingly, characteristic tissue macrophage signatures, including gene expression and epigenetic marks, are rapidly lost upon *ex vivo* culture, as best established for microglia [11,15].

Microglia have been recognized as critical players in central nervous system (CNS) development and homeostasis [16]. Specifically, microglia contribute to synaptic remodeling, neurogenesis and the routine clearance of debris and dead cells [17-21]. Microglia furthermore act as immune sensors and take part in the CNS immune defense [22]. Deficiencies affecting intrinsic microglia fitness can result in neuropsychiatric or neurologic disorders [23]. Therapeutic approaches to these 'microgliopathies' could include microglia replacement by wild type (WT) cells. Moreover, microglia replacement by BM-derived cells has also been proposed as treatment for metabolic disorders, such as adrenoleukodystrophy (ALD) and Hurler syndrome, as well as

neuroinflammatory diseases (e.g. amyotrophic lateral sclerosis, Alzheimer's) in order to slow down disease progression or improve clinical symptoms [24]. HSC gene therapy was shown to arrest the neuro-inflammatory demyelinating process in a gene therapy approach to treat metachromatic leukodystrophy (MLD) albeit with delay [25]. Of note, replacement of YS-derived microglia by HSC-derived cells is also a byproduct of therapeutic stem cell transplantations that are routinely used to treat monogenic immune disorders, such as Wiskott–Aldrich syndrome (WAS) and IL10 receptor deficiencies [26, 27]. To what extend HSC-derived cells can replace the host microglia (especially after conditioning) and if these restore functions by cross-correction remains unclear. Understanding how engrafted cells perform in the host, in particular following challenge, is therefore of considerable clinical importance, not only in HSC transplantation but also in HSC gene therapy approaches of disorders with a neurological phenotype.

Here we report a comparative analysis of YS-derived microglia and HSC graft-derived parenchymal brain macrophages. Using RNAseq and ATACseq of host and engrafted macrophages isolated from mouse BM chimeras, we show that these cells acquire microglia characteristics such as longevity, morphology and gene expression features, but still remain significantly distinct with respect to transcriptomes and epigenomes. Furthermore, host and graft cells display discrete responses to challenge by peripheral endotoxin exposure. Finally, and extending our finding to clinical settings, we confirm that in human HSC transplant patients, grafted cells also remain distinct from host microglia. Collectively, these data establish that engrafted macrophages differ from host microglia even after prolonged residence in the brain parenchyma and could have considerable clinical implications for patients treated by HSC gene therapy.

Results

BM graft-derived parenchymal brain macrophages accumulate over time post irradiation and acquire longevity and radio-resistance

Following total body irradiation (TBI), myeloid precursors enter the brain and contribute to the parenchymal macrophage compartment [28-30]. Host microglia are relatively radio-resistant and unless combined with conditioning, engraftment of the brain macrophage pool was therefore reported to be limited [29,31-33]. It furthermore remained unclear to what extent graft-derived cells acquire over time microglia characteristics, such as longevity and radio-resistance. To address this issue, we generated BM chimeras by lethally irradiating WT mice (CD45.2) (950cGy) and transplanting them with CX_3CR1^{GFP} BM (CD45.1) [34] (**Suppl. Fig. 1A**). Four months after irradiation and BM transfer (BMT), monocyte precursors in the BM and circulating blood monocytes of the chimeras were all $CD45.1^+$ GFP^+ and hence exclusively derived from the BM graft (**Suppl. Fig. 1B, C**). Analysis of the $CD45^{int}$ $CD11b^+$ $Ly6C^-$ $Ly6G^-$ microglia compartment of the chimeras revealed the presence of GFP^+ graft- and GFP^- host-derived cells (**Suppl. Fig. 1D**). In line with earlier reports, grafted cells initially constituted only a small fraction of the parenchymal brain macrophage population. However, the cells progressively replaced the host microglia (**Suppl. Fig. 1D**). The expansion of the GFP^+ infiltrate could indicate ongoing peripheral input. Alternatively, the undamaged non-irradiated CNS-resident graft-derived cells might have an advantage over the irradiated host microglia, and gradually outcompete the latter during the reported infrequent microglial local proliferation [35,36]. To distinguish between these options, we performed a tandem engraftment. Recipient mice (CD45.1/2) were irradiated twice, 15 weeks apart, followed by engraftment with CX_3CR1^{GFP} BM (CD45.1) and CX_3CR1^{Cre}:R26-$RFP^{fl/fl}$ BM (CD45.2), respectively (**Fig. 1A**). Blood analysis of the chimeras 7 weeks after the 2nd BMT, showed that monocytes were exclusive derivatives of the 2nd graft (**Fig. 1B**), as were myeloid BM precursors (**Suppl. Fig. 1E**). In contrast, analysis of the CNS compartment of the chimeras revealed the presence of three distinct macrophage populations: host microglia ($CD45.1/2^+$), cells derived from the first graft ($CD45.1^+$ GFP^+), and cells derived from the second BM graft ($CD45.2^+$ RFP^+) (**Figure 1C, D, Suppl. Fig. 1F**). Presence of cells derived from the first graft establishes that these (1) acquired radio-resistance and (2) persisted in the chimeras for more than 2 months without contribution from the periphery. Corroborating the above results, cells derived from both grafts expanded on the expense of the host microglia (**Figure 1D**). CX_3CR1^{GFP} (CD45.1) and CX_3CR1^{Cre}:R26-$RFP^{fl/fl}$ (CD45.2) cells with ramified microglia morphology were detected in the brain parenchyma (**Fig. 1E**). Collectively, these data establish that engrafted cells adopt microglia characteristics, such as relative radio-resistance, longevity and morphology.

BM cells efficiently engraft recipient brains conditioned by irradiation or myeloablation and display clonal expansion

To further characterize the engraftment process, including clonal dynamics of the HSC-derived cells, we used two complementary approaches, comprising (1) transplantation of BM isolated from *Cx3cr1*^{CreER}*:R26*^{Confetti} ('Microfetti') mice [36] (**Fig. 2 A**) and (2) transplantation of lineage negative BM cells marked by a genetic barcode prior to transplantation [37] (**Suppl. Fig. 2a**).

Cx3cr1^{CreER}*:R26*^{Confetti} BM recipients were treated at two or ten weeks post engraftment with tamoxifen (TAM) to induce stochastic expression of one of the four fluorescent reporter proteins encoded by the Confetti construct [36] (**Fig. 2 A**). Thirty weeks later, distinct brain regions of the chimeras, including the olfactory bulb, cortex, hippocampus and cerebellum, were analyzed for the presence of graft-derived labeled cells and clonal clusters (**Fig. 2 B-F**). Engraftment was seen in all brain regions analyzed (**Fig. 2 G**), although the cerebellum displayed higher frequencies of HSC-derived cells, as reported earlier [31]. Integration of engrafted cells into the endogenous microglial network was reflected by their Iba-1 expression, similar morphology and intercellular distances as compared to host microglia (**Fig. 2 C-F, Suppl. Fig. 3**). Absence of Confetti^{+} monocytes over 10 weeks after BM transplantation excluded sustained labeling of peripheral monocytes (data not shown), indicating that the Confetti-labeled macrophages observed in the 10-week TAM treatment group were derived from the initial engraftment. Single cells or clones (defined as ≥ two same-color cells with 100 µm nearest neighbor proximity) of Confetti-labeled cells were observed at similar frequency regardless of the timing of TAM treatment in the olfactory bulb, cortex and hippocampus (**Fig. 2 G-I**), suggesting comparable proliferation capacities after engraftment. The higher number of cerebellar engrafted macrophages and clones in the 10-week TAM treatment group (**Fig. 2 J**) might be attributable to niche-dependent differences in kinetics of clonal expansion [36]. Clusters of same color Confetti-labeled grafts in close proximity provide evidence of local proliferation, since slow-renewing cortical microglia in 'Microfetti' brains do not form clones over 36 weeks [36].

For the genetic barcoding procedure, lineage negative BM cells isolated from male donors (CD45.1) were transduced with lentiviral vectors harboring advanced genetic barcodes (BC32) and a BFP reporter gene [37,38]. Transduced HSC/HPCs were transferred into 8 week old female recipient mice (CD45.2) that were irradiated (950 cGy) or conditioned with the alkylsulfonate busulfan (BU) (125 mg) used for myeloablation in pediatric and adult patients [39], as well as preclinical mouse models [40,41]. Six months after transplantation, the right hemisphere of the mice was used for immunohistochemistry (eBFP, Iba-1; **Fig. 2K**). From the left hemisphere, macrophages were sorted [42] and analyzed for presence and complexity of barcodes using

bioinformatics [38]. Both conditioning protocols resulted in efficient engraftment of the periphery and the microglia compartment (**Suppl. Fig 2B,C**), as also reflected in the numbers of clones displaying unique barcodes (**Fig. 2L, M, Suppl. Fig. 2D**). Interestingly, while engrafted cells isolated from BM, blood and spleen, and also CD45^{hi} CNS cells representing hematopoietic non-microglial cell, displayed mainly shared clones, the majority of clones identified among brain macrophages of busulfan-conditioned animals did not have counterparts in other tissues and were private (**Fig. 2N).** Corroborating our above results and that of others [33,43], this suggests that a major fraction of the grafted cells originates from transduced precursors that seed the host CNS early after engraftment and is maintained by local proliferation independent from ongoing hematopoiesis. Taken together, these results establish that efficiently engrafted donor cells adopt the phenotype and distribution of resident microglia within this cellular network and expand locally with kinetics specific to brain regions and conditioning paradigms.

BM-derived parenchymal brain macrophages exhibit steady state transcriptomes distinct from host microglia

BM graft-derived parenchymal brain macrophages acquire characteristics such as ramified morphology, longevity and radio-resistance and can hence be considered engrafted microglia-like cells that could potentially be employed for therapeutic purposes. Recent studies have highlighted the impact of the tissue environment on macrophage identities, including epigenomes and expression profiles [10,11]. To test whether graft-derived microglia acquire in the CNS such a global molecular imprint, we transplanted lethally irradiated 6 week old WT mice with congenic WT BM harboring CD45.1 alleles. Nine months post transplantation, chimeras were sacrificed and brain macrophages were isolated for transcriptome and epigenome analysis by RNAseq [44] and ATACseq [45], respectively. At this time point, half of the CNS macrophages of the chimeras were of graft origin (**Fig. 3A, Suppl. Fig. 4A**).

Global RNAseq analysis of parenchymal host and graft brain macrophages isolated from individual BM chimeras revealed that engrafted cells and host microglia showed significant transcriptome overlap, clearly distinguishing them from monocytes that served as reference for a HSC-derived cell population [46] (**Fig. 3B, C**). Of the total 11,614 detected transcripts, 10,572 (91%) displayed a lower than 2 fold-difference between the engrafted cells and host microglia. On the other hand, 979 transcripts were differentially expressed between these two populations (absolute value of log2-fold change>1, p-value<0.05) (**Suppl. Fig. 5A**). Expression of 469 genes was restricted to host microglia, while 510 genes were uniquely expressed by the engrafted cells (**Suppl. Fig. 5A, B**). Engrafted macrophages, but not host microglia displayed for instance mRNA

encoding CCR2, Lysozyme, CD38, CD74, Mrc1, ApoE and Ms4a7 (**Fig. 3E, Suppl. Fig. 5B**). Differentially expressed genes included transcription factors (TFs) such as the basic helix-loop-helix TF Hes1, the Krueppel-like zinc finger TF Klf12, the retinoic acid receptor RxRg and the TGFβ-associated signal transducer Smad3, that were preferentially transcribed in engrafted cells. Conversely, engrafted macrophages displayed increased expression of the estrogen receptor Esr1, the runt-domain TF Runx3 and the macrophage survival factor Nr4a1, as compared to host cells (**Fig. 3D**). Of note, the host microglia in this case were irradiated and differences observed between the two populations could hence relate to radiation damage; transcriptomes of engrafted cells however also differed significantly from age-matched non-irradiated microglia (**Suppl. Fig. 5C**).

Supporting the notion of their microglia-like identity, engrafted cells expressed similar levels of the DNA-RNA binding protein TDP-43 encoded by the *Tardbp* gene recently implied as regulator of microglial phagocytosis [47], the Two-Pore Domain Potassium Channel THIK-1, encoded by the *Kcnk13* gene and shown to be critical for microglial ramification, surveillance, and IL-1b release [48] and the TF Mef2c, reported to restrain microglia responses [49] (**Fig. 3F**) Likewise, the graft also displayed expression of 'microglia signature' genes that have been proposed to distinguish these cells from other tissue macrophages and acute monocyte infiltrates associated with inflammation [50,51], including Fc receptor-like molecule (*Fcrls*) and TGFβ receptor (*Tgfbr*), which is critical to establish microglia identity [50] (**Fig. 3G)**. Other proposed 'microglia signature genes', such as *P2ry12*, *Tmem119, SiglecH* and *HexB* displayed either significantly reduced expression in the grafted cells or were exclusively expressed by host microglia, like the ones encoding the Sodium/glucose cotransporter 1 (*Slc2a5)*, the phosphoglycoprotein protein CD34 (*Cd34*) and the transcriptional repressors Sall1 (*Sall1*) and Sall3 (*Sall3*) (**Fig. 3H, Suppl. Fig. 5D**). Of note, lack of some microglia markers had been reported before for cells retrieved from acutely engrafted brains [8,50,50,52,53].

The expression signature of the engrafted macrophages showed a considerable overlap with the transcriptome of perivascular macrophages [54-56], including present and absent transcripts, such as *ApoE*, *Msn4a7*, *Slc2a5* and *Sall1,* respectively. Gene Set Enrichment Analysis (GSEA) [55] revealed that engrafted macrophages displayed an activation signature as compared to host microglia (**Suppl. Fig. 6A**). Finally and corroborating our data, the list of genes we report differentially expressed by engrafted and host cells also displayed a considerable overlap with results recently reported by two other groups [57,58] (**Suppl. Fig. 6B**).

Sall1, a member of the *Spalt* ('Spalt-like' (Sall)) family of evolutionarily conserved transcriptional regulators critical for organogenesis, acts as repressor by recruitment of the

Nucleosome Remodeling and Deacetylase Corepressor Complex (NurD). Binding motifs of Sall1 and hence its direct genomic targets remain undefined precluding a direct assessment of the impact of the lack of the repressor on the expression signature of the grafted macrophages. Interestingly though, comparison of the recently reported list of genes differentially expressed by WT and Sall1-deficient microglia [53] and that of host and graft brain macrophages revealed in this study, showed significant overlap (**Fig. 3H**). This included expression of genes otherwise restricted to macrophages residing in non-CNS tissues, such as *Msr1,* encoding a scavenger receptor, and *Cyp4f18* encoding cytochrome P450 (**Fig. 3I**). Furthermore, like Sall1-deficient microglia [53]**,** grafted CNS macrophages displayed an activation signature, as reflected by expression of *Cybb* encoding the Cytochrome b-245 heavy chain and *Axl* encoding a member of a tyrosine kinase receptor family critical for debris clearance [20] (**Fig. 3I, Suppl. Fig. 6B**). This suggests that a major fraction of the differential expression of host microglia and engrafted cells could be explained by the specific absence of the transcriptional repressor Sall1 from the former cells. Overall, these findings establish that engrafted macrophages that persist in the brain and acquire microglia characteristics such as morphology and radio-resistance, also show significant transcriptome overlap with host microglia, but remain a molecularly distinct population.

Engrafted CNS macrophages and host microglia exhibit distinct epigenomes

To further define engrafted cells and host microglia we performed an epigenome analysis using ATACseq that identifies open chromatin regions by virtue of their accessibility for 'tagmentation' by transposases [45]. Correlated ATACseq replicates (**Suppl Fig 7A**) performed on graft and host microglia isolated from BM chimeras detected 58,947 total accessible regions (corresponding to 16,156 genes). Corroborating the observed differential gene expression (**Fig. 3**), host microglia but not engrafted cells, displayed ATAC signals in the *Sall1* and *Klf2* loci as indicated in Integrative Genomics Viewer (IGV) tracks (**Fig. 4A**). ATAC peaks in other genomic locations, such as *ApoE* and *Ms4a7,* were restricted to genomes of engrafted cells, in line with mRNA detection in these cells, but not host microglia (**Fig. 4B**). As ATACseq does not discriminate between bound transcriptional activators and repressors, some differentially expressed loci did not show epigenetic differences. This included for instance the MHC II locus (*H2-ab1*), which displayed similar ATACseq peaks in host microglia and the graft that lack and display *H2-ab1* transcripts, respectively (**Fig. 4C**). These loci might be transcriptionally silenced, but activated upon cell stimulation. Similar ‘poised’ states, that might be revealed following challenge, can be assumed for gene loci, that displayed differential epigenomes and ATAC peaks, but were transcriptionally active in neither the host nor the engrafted cells. Finally, rare genes, such as the *Dbi* locus showed equal

expression, but differential ATAC profiles which might suggest that their transcription is driven by distinct TFs (**Fig. 4C**). Global quantification of differential ATAC peaks between the two brain macrophage populations revealed that 6 % of the total accessible regions (or 8.7 % of the associated genes) were distinct. Specifically, 1,506 peaks (corresponding to 941 genes) displayed a >4 fold significant (p-value<0.01) enrichment in host microglia and 2,176 peaks (corresponding to 1,465 genes) were increased in BM graft-derived cells (**Fig. 4D**).

To identify potential TFs responsible for the differential transcriptomes and epigenomes of engrafted and host microglia, we applied a TF Binding Analysis (TBA) machine learning model (Fonseca et al., 2018 – BioRX). TBA learns to jointly weigh 100s of motifs drawn from the JASPAR and CISBP databases to distinguish open chromatin sites from GC matched genomic background. By examining the contribution of each motif to the model's performance, TBA assigns a significance level to each motif. This analysis revealed a class of highly significant motifs correlated with open chromatin that were common to both host and graft cells and a variety of other myeloid cells (p<10e-20) such as CTCF, SPI1, and Runx binding motifs (**Suppl. Fig. 7B**). Of the more intermediately ranked motifs, 41 motifs were more important in either host or graft cells as measured by the log likelihood ratio (LL) when comparing graft and host microglia (LL>=10e-2, blue points **Fig. 4E**). Ten motifs were preferentially detected in host microglia (**Fig. 4E**), including motifs for TFs such as IRF8, which is a critical regulator of microglia identity [59], as well less characterized motifs such as the C2H2 Zinc Finger motif. The latter motif could potentially be recognized by Sall1, which contains tandem C2H2 zinc fingers. Motifs preferentially detected in grafted cells included CEBPa, RFX, and Jun motifs (**Fig. 4E**). Motifs that were preferentially detected in either graft or host cells exhibited even greater differences in comparison to other myeloid cell populations, consistent with environmental cues directing chromatin remodeling of grafted cells (**Fig. 4F**). Collectively, these observations substantiate the conclusion that while grafted cells adopt epigenetic characteristics similar to microglia, they remain distinct from host cells even after prolonged CNS residence.

HSC-derived macrophages display distinct responses compared to host microglia following peripheral endotoxin challenge

Given the significant transcriptome and epigenome differences between the host and BM-graft derived macrophages that persists for extended periods of time post-transplantation, we next examined whether the two populations are functionally distinct. To that end, chimeras were challenged nine months post transplantation by a peripheral injection of the bacterial endotoxin lipopolysaccharide (LPS), an established paradigm for inflammation associated with robust

microglia responses to systemic cytokine secretion. Host and engrafted cells were isolated from the brains of the chimeras 12 hours post LPS challenge, global RNA and ATAC sequencing were performed, and results were compared to the samples obtained from non-challenged mice presented earlier (**Fig. 3, 4**). Principle component analysis (PCA) revealed a high degree of similarity within each group, but segregation of the host and graft samples (**Fig. 5A**). Host and grafted cells responded with altered expression of 745 shared genes. 940 genes were changed in grafted cells only, and 602 genes were changed in host microglia upon LPS challenge (**Figure 5B**). Examples for these three categories are shown in **Figure 5C, D** and **Suppl. Fig. 8A, B**. Genes commonly induced by engrafted cells and host microglia in response to the LPS challenge comprised *Tnf*, *Ccl5* and *Tnfaip3*, encoding the A20 deubiquitinase that negatively regulates NF-κB–dependent gene expression. Commonly down-regulated genes comprised *Trem2*, *Cx3cr1* and *Aif1*. The graft specific response included upregulation of *Clec4e*, *Pirb*, *Isg15*, *Irf7* and down-modulation of *Mgl2*, *Cmah* and *CD36*. Genes specifically induced in host microglia encoded the scavenger receptor Marco, Gpr65, Tlr2 and Il12b. Moreover, host microglia silenced expression of *Sall1* and *Upk1b* (**Figure 5D**). Ingenuity analysis of transcriptomes of engrafted and host brain macrophages isolated from chimeras with and without peripheral LPS challenge revealed potential distinct upstream regulators acting on these populations, as well as a differential representation of activated functional pathways (**Supp. Fig. 8C**). Engrafted macrophages displayed for instance activation of pathways controlled by IL1b and Ifng and suppressed by IL-10.

ATACseq analysis revealed differential epigenome alterations between engrafted and host cells in response to the LPS challenge. Specifically, correlated ATACseq replicates (**Suppl. Fig 9A**) performed on engrafted and host cells isolated from BM chimeras after LPS challenge detected 46,485 total accessible regions (corresponding to 15,390 genes). Global quantification of differential ATAC peaks between the two brain macrophage populations revealed a total of 552 peaks (corresponding to 391 genes) that displayed a >4 fold significant (p-value<0.01) enrichment in host microglia and 841 peaks (corresponding to 618 genes) that were increased to the same extent in BM graft-derived cells (**Figure 6A**).

Overall, differences between host and graft cells were less pronounced after LPS challenge (**Figure 4D, 6A**). Analysis of motifs with TBA models trained on intergenic peaks present after LPS challenge identified a core group of highly significant motifs common to graft and host cells in both no treatment and LPS conditions that may be important for maintaining microglia identify (p<10e-20, **Suppl. Fig. 9B**). Under LPS conditions, we observed that AP-1 family motifs (Jun-related, Fos-related) and NF-kappaB (Rel, Nfkb1) motifs are more significant, which is consistent with the role AP-1 and NF-kappaB factors play in mediating the macrophage inflammatory response (**Fig. 6B**).

Interestingly, host and graft cells preferred different variants of the NF-kappaB and AP-1 motifs (motif logos, **Fig. 6B**). The NF-kappaB motif most highly enriched in host cells was more GC rich than that in graft cells, whereas the AP-1 motif most enriched in graft cells corresponded to an N(1) spaced motif (TGANTCA) in contrast to the N(2) motif (TGANNTCA) that was most highly enriched in host cells. Of the 27 motifs that were differentially detected in challenged host and graft cells, several motifs were also differentially detected in untreated cells, including motifs for IRF8, Zinc Finger factors, NK related factors, and Tal related factors (**Fig. 4F, 6B**).

Appearance of ATACseq peaks was correlated with differential gene expression between the two macrophage populations (**Figure 6C**). Transcripts encoding the scavenger receptor Marco were absent from host microglia, but specifically induced in these cells but not engrafted cells following the LPS challenge (**Figure 6C**). Likewise, the *Marco* locus (92kb) in microglia of unchallenged animals displayed 5 ATACseq peaks that were all restricted to the host cells (I-V; **Figure 6C**). LPS challenge resulted in loss of one peak (IV) and the induction of 3 additional peaks (VI- VIII), again restricted to host microglia (**Figure 6C**). An induced peak located 53,411 bp downstream of the TSS displayed a host-specific Nfkb1 motif, whereas a second peak located 19,210 bp upstream of the Marco TSS displayed a host-specific AP-1 (Fos-related) motif (**Figure 6C**). Engrafted macrophages on the other hand, displayed as compared to host microglia prominent induction of Secreted Phosphoprotein 1 (SPP1) / Osteopontin (**Figure 6D**), a factor that got recent attention as part of a microglia signature that could be associated with certain CNS pathologies [60]. In accordance with the expression results, *Spp1* loci (74kb) of engrafted macrophages displayed following LPS challenge 3 ATACseq peaks (IV- VI), that were significantly enhanced over microglia (**Figure 6D**). An induced peak located 54,402 bp upstream of the TSS displayed a TBX5 motif, whereas a second peak located 38,600 bp upstream of the Spp1 TSS displayed an AP-1 (Jun-related) motif (**Figure 6D**). Collectively these data establish that engrafted microglia respond differently from host microglia to a challenge and are hence functionally distinct.

HSC-derived macrophages in murine and human brain parenchyma differ from host microglia

Engrafted brain macrophages differ from host microglia by their gene expression (**Fig. 3B**). To confirm this finding for protein expression we performed a histological analysis of brains of the BM chimeras generated by TBI and BU conditioning (**Suppl. Figure 2A**). Engrafted cells and host microglia were identified by Iba-1 staining. Graft-derived cells were defined according to eBFP expression conveyed by the lentiviral construct (**Figure 2K**). In concordance with the transcriptome

data (**Figure 3G**), analysis for Tmem119 and P2ry12 expression revealed absence of these markers from the graft (**Fig. 7 A, B**), while eBFP$^+$ cells displayed ApoE and MHC Class II (**Suppl. Fig. 10**).

To finally extrapolate our finding to a human setting we analyzed *post mortem* brains of patients that had underwent HSC transplantation. Specifically, we took advantage of gender-mismatched grafts that allowed identification of the transplant by virtue of its Y-chromosomes through chromogenic *in situ* hybridization (CISH). Ramified Iba1$^+$ microglia-like cells harboring the Y chromosome could be readily identified juxtaposed to Y chromosome-negative host microglia in cortex, cerebellum and hippocampus sections of the subject brains (**Fig. 7 C**). Expression of the purinergic $P2Y_{12}$ receptor has been proposed earlier to serve as marker for human microglia [52, 61]. Moreover, absence of $P2Y_{12}$ compromises microglial activation by nucleotides and could thus have functional implications [62]. As observed in the murine chimeras, also human $P2Y_{12}$ receptor expression was found to be restricted to host cells, but absent from Y chromosome-positive brain macrophages cells (**Fig. 7 D**). Collectively these data establish that in the CNS of patients that underwent HSC-transplantation graft-derived cells remained functionally distinct from host microglia and strengthen the conclusion that HSC-derived engrafted cells differ from host microglia.

Discussion

Here we established that HSC-derived brain macrophages that persistently seed the CNS of recipient organisms following irradiation or myelo-ablation remain distinct from host microglia with respect to their transcriptomes, epigenomes and response to challenge.

Following the engraftment of conditioned recipient mice, transplanted cells establish under the influence of the CNS microenvironment a characteristic microglia transcriptome that distinguishes these cells from other tissue macrophages [50]. Thus, nine tenth of their transcriptome is shared with host microglia, including expression of Fc receptor-like molecules (*Fcrls*) and Tgfb receptor (*Tgfbr*), as well as the MADS box transcription enhancer factor 2 (*Mef2c*). Moreover, residence in the CNS endowed engrafted cells with microglia characteristics, such as longevity, radio-resistance and ramified morphology. Engrafted cells however failed to adopt complete host microglia identity even after prolonged CNS residence. Corroborating and extending earlier reports [50,52,53], this included significantly reduced mRNA expression of the microglia markers *Tmem119*, *SiglecH* and *P2yr12* and complete lack of the transcriptional repressors Sall1 and Sall3. Conversely, engrafted macrophages expressed genes absent from host microglia, including *Ccr2*, *Ifnar1*, *Msa4a7* and *ApoE*, and displayed *Msr1* and *Axl* mRNAs, potentially related to the absence

of Sall1 [53]. Transcriptomes of engrafted macrophages showed considerable overlap with perivascular macrophages and indication of cell activation, such as an underrepresentation of a regulatory pathway driven by IL-10, as compared to host microglia. Comparative transcriptome analysis showed that our data are well in line with recent studies [57,58] that reported that macrophages that engraft the brain of mice conditionally depleted of microglia due to a Csf1r deficiency also retain a transcriptional identity distinct from host cells.

Analysis of open chromatin revealed that graft cells acquire a transposase-accessible profile that is similar to that of host cells and is enriched for a common set of motifs that are recognized by TFs known to be important for microglia development, such as PU.1 and MEF2c. However, host and graft cells also exhibit significant differences in open chromatin that are associated with distinct motif enrichment patterns and the observed differences in gene expression. These findings imply that the distinct developmental origins of host and graft cells determine the ability of the brain environment to fully activate the complement of transcription factors required for microglia identity, most notably exemplified by lack of induction of *Sall1* in graft cells. The differences in chromatin landscapes under resting conditions are likely to contribute to the host and graft-specific responses to LPS challenge. This possibility is supported by the observation that alternative motifs for NFkB and AP-1 factors are enriched in open chromatin of host and graft cells. We interpret this finding to reflect the binding of NFkB and AP-1 dimers and heterodimers to different locations in the genome that are specified by the specified by host or graft-specific combinations of transcription factors, respectively.

The exact origin of the HSC-derived engrafted cells in the chimeric organisms remains to be defined. In a classic study, Ajami and colleagues established that non-parenchymal brain macrophages that can persistently seed the host brain originate from non-monocytic BM-resident myeloid progenitors characterized by the absence of CX_3CR1 expression [28]. Likewise, studies by Biffi and co-workers suggested that a transient wave of early hematopoietic progenitors infiltrates the host CNS during transplantation and following local proliferation, establish the graft [33]. This notion is supported by the results of our 'Microfetti' and barcoding approaches that establish that engrafted macrophages undergo clonal proliferation and thereby likely progressively outcompete irradiation or busulfan-damaged host microglia. Moreover, the conclusion that engrafted cells arose from cells that do not contribute to long-term hematopoiesis in the chimeras is also in line with the prominent detection of private clones in this population, which are not shared with the other hematopoietic compartments.

Future studies could aim to identify cells that upon engraftment will give rise to closer mimics of host microglia, including for instance expression of Sall1. This could include cells linked

with the unique developmental YS origin of microglia [57], or otherwise manipulated cells such as microglia-like cells derived from induced-pluripotent-stem-cell (iPS)-derived primitive macrophages [63]. Furthermore, in the context of gene therapy, viral vectors could be used to express transgenes to engineer the engrafted cells to boost engraftment and modulate their function. Of note however, under certain pathological conditions the distinct engrafted HSC-derived macrophages we report might also be advantageous as compared to host microglia. HSC-derived cells could for instance be superior to YS-derived microglia in the handling of the debris burden associated with senescence [64] or amyloid plaques that arise during Alzheimer's disease [65]. Indeed, the latter hypothesis was proposed although the exact underlying mechanism remains unclear [66]; elucidation of such scenarios should profit from the molecular definition of the engrafted cells and host microglia provided in this study. However, *in vivo* functions of microglia remain poorly understood and future dedicated experimentation will be required to compare the performance of engrafted macrophages and host microglia in different disease models during aging and specific challenges.

Recent studies revealed a signature of disease associated microglia, termed 'DAM' [55] or 'MGnD' [67], that is induced by various brain-intrinsic changes in the absence of massive peripheral infiltrates, though not following peripheral LPS challenge [60,68], and can occur on acute to chronic time scales [60]. Of note, engrafted brain macrophages displayed robust constitutive expression of some of theses DAM genes, such as *ApoE* and *Axl*. Likewise, as opposed to microglia [60], grafted HSC-derived brain macrophages responded to the LPS challenge with the induction of DAM/ MGnD hallmarks, such as the CD44 ligand Spp1/Osteopontin. The latter might have implications when considering brain macrophage contributions to CNS pathologies. Results obtained from fate mapping models currently suggest that at least in the brain of unchallenged C57BL6 mice kept in specific-pathogen-free (SPF) facilities, parenchymal macrophages are exclusively comprised of YS-derived derived microglia. Further experimentation will however be required to assess how absolute this exclusion of HSC-derived macrophages is, in particular following challenges. Moreover, it remains currently unclear to what extend HSC-derived macrophages might be able to seed the human brain, f.i. during extended aging, and could hence impact brain pathologies.

Tissue macrophages such as Kupffer cells (KC) and alveolar macrophages (AM) have been reported to be faithfully replaced by HSC-derived cells in irradiation chimeras and other small animal models involving deficiencies of the resident compartment [6-9]. While these studies were restricted to transcriptome comparison and hence might have missed epigenetic differences between graft and host cells, the inability of HSC-derived cells to achieve full host cell identity might be unique to microglia and related to features particular to these cells. Specifically, among

adult tissue macrophages, only microglia derive from primitive YS macrophages and this origin could define cell identity. In contrast, generation of both KC and AM involves a monocytic intermediate, and their re-generation might hence be attainable by the closer related HSC-derived cells that can also give rise to monocytes. Alternatively, establishment of the '*bona fide*' microglia signature might require 'physiological' microglia development in the developing CNS that is associated with profound transient activation of this brain macrophage compartment [12,13,59,69].

Collectively, the demonstration that engrafted cells and host microglia remain distinct sheds light on the molecular and functional heterogeneity of parenchymal brain macrophages. Moreover, when extrapolated to the human setting, our findings could have major implications for patients treated by HSC gene therapy to ameliorate lysosomal storage disorders, microgliopathies or general monogenic immune-deficiencies.

Materials and Methods

Mice

For generation of BM chimeras, wild type C57BL/6 J mice (Harlan) were used as recipients, $Cx3cr1^{GFP}$ or CX_3CR1^{Cre}:R26-$RFP^{fl/fl}$ mice [34,70] were used as donors. Recipient mice were lethally irradiated with a single dose of 950 cGy using an XRAD 320 machine (Precision X-Ray (PXI)) and reconstituted the next day by i.v. injection of 5 x 10^6 donor BM cells per mouse. All animals bred at the Weizmann animal facility were maintained under specific pathogen-free conditions and handled according to protocols approved by the Weizmann Institute Animal Care Committee as per international guidelines. Female C57BL/6 J wild type recipient mice (Charles River) received full body irradiation (7 Gy) in the RS 2000 Biological Research Irradiator at 8 weeks of age. After 24 h, 5 x 10^6 donor BM cells isolated from $Cx3cr1^{CreER/+}$:$R26R^{Confetti/+}$ 'Microfetti' mice [38] were injected into the tail vein of recipients. Recipient mice received 500 µl Cotrim K-ratiopharm® antibiotics in 250 ml drinking water for two weeks after BM reconstitution. Tamoxifen (Sigma) dissolved in corn oil (Sigma) was subcutaneously applied in a single dose of 10 mg distributed along the flanks. Mice were maintained in specific-pathogen-free facility with chow and water provided ad libitum. Animal experiments performed in Freiburg were approved by the Regional Council of Freiburg, Germany. All experimenters were blinded during data acquisition and analysis.

Barcoding experiment

Female 8 weeks-old C57Bl6J wild type recipient mice (CD45.2) were either conditioned with total body irradiation (9.5 Gy) or 125mg Busulfan per kg bodyweight (in 5 doses à 25mg) (Wilkinson 2012) prior to transplantation of 5 x 10^5 lineage negative cells from male CD45.1-mice (8 weeks of

age), transduced with lentiviral BC32-eBFP vectors [37]. Intermediate peripheral blood samples were taken every 4-6 weeks to monitor chimerism and marking efficiency via flow cytometry. Six months post transplantation, the mice were sacrificed and peripheral blood, bone marrow, spleen and the brain were taken for clonal analyses. Mice were maintained in specific-pathogen-free facility with chow and water provided ad libitum. Animal experiments performed in Hamburg were approved by the local authorities (Behoerde fuer Gesundheit und Verbraucherschutz-Veterinaerwesen/Lebensmittelsicherheit).

LPS challenge

For LPS treatment, mice were injected intra-peritoneally (i.p.) with a single dose of lipopolysaccharide (LPS) (2.5 mg/kg, *E. coli* 0111:B4; Sigma) and sacrificed 12 hours post-injection.

Microglia isolation

To isolate microglia and HSC-derived parenchymal CNS macrophages, BM chimeras were perfused using ice-cold phosphate buffered saline (PBS) and brains were harvested. Brains were dissected, homogenized by pipetting and incubated for 20 min at 37°C in a 1 ml HBSS solution containing 2% BSA, 1 mg/ml Collagenase D (Sigma) and 1 mg/ml DNase1 (Sigma). The homogenate was then filtered through a 100 µm mesh and centrifuged at 2200 RPM, at 4°C, for 5 min. For the enrichment of microglia and BM-derived cells, the pellet was re-suspended with a 40% percoll solution (Sigma) and centrifuged at 2200 RPM, room temperature for 15 min. The cell pellet was next subjected to antibody labeling and flow-cytometry analysis.

Flow cytometry and cell sorting

Cells were stained with primary antibodies against CD45.1 (A20), CD45.2 (104), CD11b (M1/70), Ly6C (AL-21), and LY6G (1A8) - all from Biolegend, San Diego, CA, USA. After incubation with the Abs at 4°C for 15 min, cells were washed and sorted using a FACSAria (BD, Erembodegem, Belgium) flow cytometer. Data were acquired with FACSdiva software (Becton Dickinson). Post-acquisition analysis was performed using FlowJo software (Tree Star, FlowJo LLC; Ashland, Oregon).

RNA-seq

RNA-seq of populations was performed as described previously [10]. In brief, 10^3 - 10^5 cells from each population were sorted into 50 µL of lysis/binding buffer (Life Technologies) and stored at 80

C. mRNA was captured with Dynabeads oligo(dT) (Life Technologies) according to manufacturer's guidelines. We used a derivation of MARS-seq [44] to prepare libraries for RNA-seq. Briefly, RNA was reversed transcribed with MARS-seq barcoded RT primer in a 10 µL volume with the Affinity Script kit (Agilent). Reverse transcription was analyzed by qRT-PCR and samples with a similar CT were pooled (up to 8 samples per pool). Each pool was treated with Exonuclease I (NEB) for 30 min at 37 C and subsequently cleaned by 1.2X volumes of SPRI beads (Beckman Coulter). Subsequently, the cDNA was converted to double-stranded DNA with a second strand synthesis kit (NEB) in a 20 mL reaction, incubating for 2 hr at 16 C. The product was purified with 1.4X volumes of SPRI beads, eluted in 8 µL and in vitro transcribed (with the beads) at 37 C overnight for linear amplification using the T7 High Yield RNA polymerase IVT kit (NEB). Following IVT, the DNA template was removed with Turbo DNase I (Ambion) 15 min at 37 C and the amplified RNA (aRNA) purified with 1.2 volumes of SPRI beads. The aRNA was fragmented by incubating 3 min at 70 C in Zn^{2+} RNA fragmentation reagents (Ambion) and purified with 2X volumes of SPRI beads. The aRNA was ligated to the MARS-seq ligation adaptor with T4 RNA Ligase I (NEB)." The reaction was incubated at 22 C for 2 hr. After 1.5X SPRI cleanup, the ligated product was reverse transcribed using Affinity Script RT enzyme (Agilent) and a primer complementary to the ligated adaptor. The reaction was incubated for 2 min at 42 C, 45 min at 50 C, and 5 min at 85 C. The cDNA was purified with 1.5X volumes of SPRI beads. The library was completed and amplified through a nested PCR reaction with 0.5 mM of P5_Rd1 and P7_Rd2 primers and PCR ready mix (Kappa Biosystems). The amplified pooled library was purified with 0.7X volumes of SPRI beads to remove primer leftovers. Library concentration was measured with a Qubit fluorometer (Life Technologies) and mean molecule size was determined with a 2200 TapeStation instrument. RNA-seq libraries were sequenced using the Illumina NextSeq 500. Raw reads were mapped to the genome (NCBI37/mm9) using hisat (version 0.1.6). Only reads with unique mapping were considered for further analysis. Gene expression levels were calculated using the HOMER software package (analyzeRepeats.pl rna mm9 -d < tagDir > -count exons -condenseGenes -strand + -raw) (Heinz et al., 2010). Normalization and differ- ential expression analysis was done using the DESeq2 R-package. Differential expressed genes were selected using a 2-fold change cutoff between at least two populations and adjusted p value for multiple gene testing > 0.05. Gene expression matrix was clustered using k-means algorithm (MATLAB function kmeans) with correlation as the distance metric. The value of k was chosen by assessing the average silhouette (MATLAB function silhouette) (3) for a range of possible values (4–15).

ATAC-seq

20,000-50,000 cells were used for ATAC-seq [45] applying described changes (Lara-Astiaso et al., 2014). Briefly, nuclei were obtained by lysing the cells with cold lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Igepal CA-630) and nuclei were pelleted by centrifugation for 20 min at 500 g, 4 C using a swing rotor. Supernatant was discarded and nuclei were re-suspended in 25 µL reaction buffer containing 2 µL of Tn5 transposase and 12.5 µL of TD buffer (Nextera Sample preparation kit from Illumina). The reaction was incubated at 37 C for 1 hr. DNA was released from chromatin by adding 5 µL of clean up buffer (900 mM NaCl, 300 mM EDTA, 1.1% SDS, 4.4 mg/ml Proteinase K (NEB)) followed by an incubation for 30 min at 40 C. Tagmentated DNA was isolated using 2X volumes of SPRI beads and eluted in 21 µl. For library amplification, two sequential PCRs (9 cycles, followed by an additional 6 cycles) were performed in order to enrich small tagmentated DNA fragments. We used the indexing primers as described by Buenrostro et al., 2013 and KAPA HiFi HotStart ready mix. After the first PCR, the libraries were size-selected using double SPRI bead selection (0.65X followed by 1.8X). Then the second PCR was performed with the same conditions in order to obtain the final library. DNA concentration was measured with a Qubit fluorometer (Life Technologies) and library sizes were determined using TapeStation (Agilent Technologies). Libraries where sequenced on the Illumina NextSeq 500 obtaining an average of 20 million reads per sample. Putative open chromatin regions (peaks) were called using HOMER [71] (using parameters compatible with IDR analysis: –L 0 –C 0 –fdr 0.9). The Irreproducible Discovery Rate (IDR) was computed for each peak using the Homer peak score for each replicate experiment (https://github.com/nboley/idr); peaks with IDR > 0.05 were filtered away. Intergenic peaks annotated by Homer were used to train TBA models for each cell type using default parameters (https://github.com/jenhantao/tba). The significance of each motif in each cell type was assigned by comparing the predictive performance of the trained TBA model and a perturbed model that cannot recognize one motif using the chi-squared test.

Motif Analysis with TBA

We used TBA (a Transcription factor Binding Analysis) models to identify enriched motifs in open chromatin in comparison to a set of randomly selected genomic loci (matched for GC content). For each of the sequences in the combined set of the ATAC-seq peaks and background loci, TBA calculates the highest motif score for each of the motifs (in either orientation) included in the model. The sequences of open chromatin regions and background loci as well as the corresponding motif scores are used to train the TBA model to distinguish open chromatin from background loci (using five-fold cross validation). A TBA model scores the probability of observing open chromatin given a genomic sequence by computing a weighted sum over all the motif scores

computed for that sequence. The weight for each motif is learned by iteratively modifying the weights (starting from random values) until the model's ability to differentiate open chromatin from background sequences no longer improves. The significance of a given motif was assigned by comparing the predictive performance of a trained TBA model and a perturbed model that cannot recognize that motif using the likelihood ratio test. We trained TBA models for each cell type, using version 1.0 of TBA and default parameters (source code and executable files are available at: https://github.com/jenhantao/tba). For a complete description see Fonseca et al., 2018 {bioRx deposition}

Histology

Mice were transcardially perfused with PBS followed by 4% paraformaldehyde in PBS. Mouse brains were post-fixed for 6 h or overnight at 4°C and processed for frozen sectioning as previously detailed [36]. Free floating 50-µm cryosections and 14-µm cryosections on slides were prepared from Microfetti and lineage negative barcoded BM chimeras, respectively. Tissues were permeabilized in blocking solution (0.1% Triton-X 100, 5% bovine albumin, and PBS) for 2 h at room temperature and incubated overnight at 4°C with primary antibodies: 1:500 rabbit anti-Iba-1 (Wako), 1:200 goat anti-Iba-1 (Novus), 1:1000 chicken anti-GFP against eBFP (Abcam), 1:1000 rabbit anti-TMEM119 (Abcam), 1:500 rabbit anti-P2RY12 (Ana Spec), 1:200 goat anti-APOE (Millipore), and 1:100 mouse anti-MHC Class II (Abcam). Antigen retrieval was performed prior to APOE staining for 40 min at 92 °C in pH 9 citrate buffer. Corresponding secondary antibodies conjugated to Alexa Fluor 488, Alexa Fluor 568, or Alexa Fluor 647 (Life Technologies) were applied at 1:1000 with nuclear counterstain by 4',6-diamidino-2-phenylindole (DAPI, Sigma) at 1:5000 for 2 h at room temperature. Sections were mounted in ProLong® Diamond Antifade Mountant (Life Technologies).

Image acquisition and analysis

Brain images of Microfetti and lineage negative BM chimeras were acquired on the Keyence BZ-9000 inverted fluorescence microscope using a 20X / 0.75 NA objective lens. Images were processed and analyzed using ImageJ (NIH). Intercellular distances were determined by CalculateNearestNeighbor from Kota Miura (http://doi.org/10.5281/zenodo.1323726). Cell morphological analyses were performed with ImageJ plugins Simple Neurite Tracer [72] and Sholl Analysis [73] using default settings.

Human gender-mismatched stem cell transplantation patients

Experiments on human tissue samples were performed according to the Declaration of Helsinki. Ethical approval was obtained from the local Research Ethics Committee of the University Medical Center of Freiburg (ref. no. 10008/09). The analyzed brain tissue samples of female patients, who underwent gender-mismatched peripheral blood stem cell transplantation (PBSCT), were derived from the brain autopsy case archive of the Institute of Neuropathology, University of Freiburg. Patient 1 was diagnosed with myelo-dysplastic syndrome (MDS) 12 years before death and received a first sex-matched PBSCT six years later. Graft failure developed a year later. The patient's MDS transformed into acute myeloid leukemia (AML) 17 months prior to death. Patient 1 then received a second PBSCT from a human leukocyte antigen (HLA)-identical male donor. Transplantation was initiated after a myelo-ablative therapy regimen comprising Thiotepa, Fludarabine, and Treosulfan. Complications after transplantation included mucositis and cytomegalovirus reactivation. Complete chimerism was confirmed in two follow-up examinations. Patient 1 died at the age of 66 years 453 days after sex-mismatched PBSCT due to pneumonia and pulmonary leukostasis caused by AML infiltration. Patient 2 was diagnosed with multiple myeloma 4.5 years before death and first treated with several chemotherapy regimens. She received PBSCT from a HLA-identical male donor 1.5 years after diagnosis. Myeloablation was performed with Thiotepa, Fludarabine, and Busulfan. After transplantation, Patient 2 developed graft-versus-host disease of the skin and gut. Follow-up examinations confirmed complete chimerism. Patient 2 died at the age of 58 years 1018 days after sex-mismatched PBSCT due to pneumonia causing respiratory failure.

Combined immunohistochemistry and *in situ* hybridization

After the death of the two patients, brains were transferred into 4% paraformaldehyde (PFA) within 48 h and fixed for at least 3 weeks. After fixation, representative tissues from several brain regions including the temporal cortex, hippocampus, and cerebellum were dissected and embedded in paraffin. Routine neuropathological examination revealed minimal signs of neurodegeneration in Patient 1, according to Braak stage I-II [74]. Patient 2 showed no signs of brain pathology. Combined immunohistochemistry (IHC) and chromogenic in situ hybridization (CISH) were performed on 10 µm thick sections. Sections were deparaffinized and heated at 92 °C, at pH 6 for 40 min for antigen retrieval. For IHC, the sections were incubated with the respective primary antibodies for 30 min at room temperature: 1:1000 rabbit anti-IBA-1 (Abcam, clone EPR 16588), 1:1500 rabbit anti-P2RY12 (Sigma-Aldrich, polyclonal). Secondary goat anti-rabbit antibodies (Southern Biotech) were applied at 1:200 for 1 h at room temperature. Liquid Permanent Red Substrate-Chromogen (Agilent Dako) was used for visualization of the antigen. After 5 min post-fixation in 4 % PFA and rinse in

deionized water, CISH was performed on the sections using the ZytoDot CISH Implementation Kit (ZytoVision) according to manufacturer's instructions with the following modifications. After incubation in ethylene diamine tetra-acetic acid (EDTA) at 95 °C for 15 min, sections were treated with pepsin solution at 37°C for 6 min. After stepwise dehydration, 12 µl of ZytoDot CEN Yq12 digoxigenin-linked probe (ZytoVision) was added to each brain section to bind to the Yq12 region of the human Y-chromosome for at least 20 h at 37 °C. Following the washing and blocking steps, the sections were incubated in mouse anti-digoxigenin antibody solution, treated with horseradish peroxidase-conjugated anti-mouse antibody at 37 °C for 30 min, and bound to 3,3′-diaminobenzidine at 37 °C for 45 min. Nuclei were counterstained with hematoxylin. Images were acquired using MikroCam II with a UPlan FLN 40x / 0.75 NA objective on a BX40 microscope (Olympus).

Barcode analyses

DNA was extracted from peripheral blood, bone marrow, spleen and sorted microglia and $CD45^{high}$ cells were used for barcode amplification, multiplexing, as well as the bioinformatic processing [39] . Unique barcodes with more than 100 reads per sample were taken into further analyses. Venn diagrams were produced by an in-house R-script using the "VennDiagram" package.

Digital droplet PCR

To determine the chimerism in the sorted microglia, digital droplet PCR was performed. In a duplex reaction, a Y-chromosome-specific fragment (and a control amplicon (located in the erythropoietin receptor) were simultaneously amplified and analysed using the QX200 system (BioRad).

Statistical analysis

Mean data are shown. Mann-Whitney test, D'Agostino-Pearson test, two-tailed unpaired *t*-test, and Welch's *t*-test were performed in GraphPad Prism7. Statistical significance was taken at $P < 0.05$.

Author contributions

A. S., T.L.T, A.V. J.-S.K., M.G., P. S. performed experiments; J. G., J.T., E.D. and L.T. performed bioinformatics analysis, A. A.-F.and L.C.-M. provided technical help; C. K. G., K. C., M. P. supervised and advised; S.J. wrote the manuscript, secured funding and and supervised.

Acknowledgements

The Jung laboratory was supported by the Israeli Science Foundation (887/11), the European Research Council (Adv ERC 340345), the Deutsche Forschungsgemeinschaft (CRC/TRR167 'NeuroMac'), the U.S.-Israel Binational Science Foundation (BSF) and a collaborative network grant of the International Progressive MS Alliance (PMSA). M.P. is supported by the BMBF-funded competence network of multiple sclerosis (KKNMS), the Sobek Foundation, the Ernst-Jung Foundation, the DFG (SFB 992, SFB1160, SFB/TRR167, Reinhart-Koselleck-Grant) and the Ministry of Science, Research and Arts, Baden-Wuerttemberg (Sonderlinie "Neuroinflammation"). TLT was supported by the German Research Foundation (DFG, TA1029/1-1) and Ministry of Science, Research and the Arts of Baden-Württemberg (7532.21/2.1.6). The Glass laboratory was support by NIH grants NS096170, DK091183 and GM085764. The authors thank G. Friedlander for help with bioinformatics, T. Sonntag, E. Orthey and the UKE FACS Core Facility from the University Medical Center Hamburg-Eppendorf, as well as J. Dautzenberg and E. Barleon from the University of Freiburg for excellent technical assistance.

References

1. Stremmel, C. *et al.* Yolk sac macrophage progenitors traffic to the embryo during defined stages of development. *Nat Comms* 1–14 (2017). doi:10.1038/s41467-017-02492-2
2. Ginhoux, F. *et al.* Supporting Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* **330,** 841–845 (2010).
3. Schulz, C. *et al.* A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* **336,** 86–90 (2012).
4. Varol, C., Mildner, A. & Jung, S. Macrophages: development and tissue specialization. *Annu. Rev. Immunol.* **33,** 643–675 (2015).
5. Guilliams, M. & Scott, C. L. Does niche competition determine the origin of tissue-resident macrophages? *Nature Publishing Group* **17,** 451–460 (2017).
6. Scott, C. L. *et al.* Bone marrow-derived monocytes give rise to self-renewing and fully differentiated Kupffer cells. *Nat Comms* **7,** 10321 (2016).
7. Gibbings, S. L. *et al.* Transcriptome analysis highlights the conserved difference between embryonic and postnatal-derived alveolar macrophages. **126,** 1357–1366 (2015).
8. Bruttger, J. *et al.* Genetic Cell Ablation Reveals Clusters of Local Self- Renewing Microglia in the Mammalian Central Nervous System. *Immunity* **43,** 92–106 (2015).
9. van de Laar, L. *et al.* Yolk Sac Macrophages, Fetal Liver, and Adult Monocytes Can Colonize an Empty Niche and Develop into Functional Tissue-Resident Macrophages. *Immunity* 1–15 (2016). doi:10.1016/j.immuni.2016.02.017
10. Lavin, Y. *et al.* Tissue-Resident Macrophage Enhancer Landscapes Are Shaped by the Local Microenvironment. *Cell* **159,** 1312–1326 (2014).
11. Gosselin, D. *et al.* Environment Drives Selection and Function of Enhancers Controlling Tissue-Specific Macrophage Identities. *Cell* **159,** 1327–1340 (2014).

12. Matcovitch-Natan, O. *et al.* Microglia development follows a stepwise program to regulate brain homeostasis. *Science* **353,** aad8670 (2016).

13. Mass, E. *et al.* Specification of tissue-resident macrophages during organogenesis. *Science* **353,** aaf4238–aaf4238 (2016).

14. T'Jonck, W., Guilliams, M. & Bonnardel, J. Niche signals and transcription factors involved in tissue-resident macrophage development. *Cell. Immunol.* (2018). doi:10.1016/j.cellimm.2018.02.005

15. Bohlen, C. J. *et al.* Diverse Requirements for Microglial Survival, Specification, and Function Revealed by Defined-Medium Cultures. *Neuron* **94,** 759–773.e8 (2017).

16. Li, Q. & Ben A Barres. Microglia and macrophages in brain homeostasis and disease. *Nature Publishing Group* 1–18 (2017). doi:10.1038/nri.2017.125

17. Sierra, A. *et al.* Microglia Shape Adult Hippocampal Neurogenesis through Apoptosis-Coupled Phagocytosis. *Stem Cell* **7,** 483–495 (2010).

18. Squarzoni, P. *et al.* Microglia modulate wiring of the embryonic forebrain. *CellReports* **8,** 1271–1279 (2014).

19. Schafer, D. P. *et al.* Microglia Sculpt Postnatal Neural Circuits in an Activity and Complement-Dependent Manner. *Neuron* **74,** 691–705 (2012).

20. Fourgeaud, L. *et al.* TAM receptors regulate multiple features of microglial physiology. *Nature* **532,** 240–244 (2016).

21. Tremblay, M.-È., Lowery, R. L. & Majewska, A. K. Microglial interactions with synapses are modulated by visual experience. *Plos Biol* **8,** e1000527 (2010).

22. Klein, R. S. & Hunter, C. A. Protective and Pathological Immunity during Central Nervous System Infections. *Immunity* **46,** 891–909 (2017).

23. Prinz, M. & Priller, J. Microglia and brain macrophages in the molecular age: from origin to neuropsychiatric disease. *Nat Rev Neurosci* **15,** 300–312 (2014).

24. Biffi, A. Hematopoietic Gene Therapies for Metabolic and Neurologic Diseases. *Hematol. Oncol. Clin. North Am.* **31,** 869–881 (2017).

25. Biffi, A. *et al.* Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* **341,** 1233158–1233158 (2013).

26. Filipovich, A. H. *et al.* Impact of donor type on outcome of bone marrow transplantation for Wiskott-Aldrich syndrome: collaborative study of the International Bone Marrow Transplant Registry and the National Marrow Donor Program. *Blood* **97,** 1598–1603 (2001).

27. Glocker, E.-O. *et al.* Inflammatory Bowel Disease and Mutations Affecting the Interleukin-10 Receptor. *N. Engl. J. Med.* **361,** 2033–2045 (2009).

28. Ajami, B., Bennett, J. L., Krieger, C., Tetzlaff, W. & Rossi, F. M. V. Local self-renewal can sustain CNS microglia maintenance and function throughout adult life. *Nat Neurosci* **10,** 1538–1543 (2007).

29. Hashimoto, D. *et al.* Tissue-Resident Macrophages Self-Maintain Locally throughout Adult Life with Minimal Contribution from Circulating Monocytes. *Immunity* **38,** 792–804 (2013).

30. Mildner, A. *et al.* Microglia in the adult brain arise from Ly-6ChiCCR2+ monocytes only under defined host conditions. *Nat Neurosci* **10,** 1544–1553 (2007).

31. Priller, J. *et al.* Targeting gene-modified hematopoietic cells to the central nervous system:

use of green fluorescent protein uncovers microglial engraftment. *Nat Med* **7,** 1356–1361 (2001).

32. Lassmann, H., Schmied, M., Vass, K. & Hickey, W. F. Bone marrow derived elements and resident microglia in brain inflammation. *Glia* **7,** 19–24 (1993).

33. Capotondo, A. *et al.* Brain conditioning is instrumental for successful microglia reconstitution following hematopoietic stem cell transplantation. *Proceedings of the National Academy of Sciences* **109,** 15018–15023 (2012).

34. Jung, S. *et al.* Analysis of fractalkine receptor CX(3)CR1 function by targeted deletion and green fluorescent protein reporter gene insertion. *Molecular and Cellular Biology* **20,** 4106–4114 (2000).

35. Askew, K. *et al.* Coupled Proliferation and Apoptosis Maintain the Rapid Turnover of Microglia in the Adult Brain. *CellReports* **18,** 391–405 (2017).

36. Tay, T. L. *et al.* A new fate mapping system reveals context-dependent random or clonal expansion of microglia. *Nat Neurosci* **20,** 793–803 (2017).

37. Aranyossy, T., Thielecke, L., Glauche, I., Fehse, B. & Cornils, K. Genetic Barcodes Facilitate Competitive Clonal Analyses In Vivo. *Human Gene Therapy* **28,** 926–937 (2017).

38. Thielecke, L. *et al.* Limitations and challenges of genetic barcode quantification. *Sci. Rep.* **7,** 43249 (2017).

39. Santos, G. W. *et al.* Marrow transplantation for acute nonlymphocytic leukemia after treatment with busulfan and cyclophosphamide. *N. Engl. J. Med.* **309,** 1347–1353 (1983).

40. Wilkinson, F. L. *et al.* Busulfan Conditioning Enhances Engraftment of Hematopoietic Donor-derived Cells in the Brain Compared With Irradiation. *Molecular Therapy* **21,** 868–876 (2013).

41. Kierdorf, K., Katzmarski, N., Haas, C. A. & Prinz, M. Bone marrow cell recruitment to the brain in the absence of irradiation or parabiosis bias. *PLoS ONE* **8,** e58544 (2013).

42. Gelderblom, M. *et al.* IL-23 (Interleukin-23)-Producing Conventional Dendritic Cells Control the Detrimental IL-17 (Interleukin-17) Response in Stroke. *Stroke* **49,** 155–164 (2018).

43. Ajami, B., Bennett, J. L., Krieger, C., Mcnagny, K. M. & Rossi, F. M. V. Infiltrating monocytes trigger EAE progression, but do not contribute to the resident microglia pool. *Nat Neurosci* **14,** 1142–1149 (2011).

44. Jaitin, D. A. *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343,** 776–779 (2014).

45. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Meth* **10,** 1213–1218 (2013).

46. Mildner, A. *et al.* Genomic Characterization of Murine Monocytes Reveals C/EBPβ Transcription Factor Dependence of Ly6C- Cells. *Immunity* **46,** 849–862.e7 (2017).

47. Paolicelli, R. C. *et al.* TDP-43 Depletion in Microglia Promotes Amyloid Clearance but Also Induces Synapse Loss. *Neuron* **95,** 297–308.e6 (2017).

48. Madry, C. *et al.* Microglial Ramification, Surveillance, and Interleukin-1β Release Are Regulated by the Two-Pore Domain K+ Channel THIK-1. *Neuron* **97,** 299–312.e6 (2018).

49. Deczkowska, A. *et al.* Mef2C restrains microglial inflammatory response and is lost in brain

ageing in an IFN-I-dependent manner. *Nat Comms* **8,** 717 (2017).

50. Butovsky, O. *et al.* Identification of a unique TGF-β-dependent molecular and functional signature in microglia. *Nat Neurosci* **17,** 131–143 (2014).

51. Gosselin, D. *et al.* An environment-dependent transcriptional network specifies human microglia identity. *Science* **1617,** eaal3222–18 (2017).

52. Bennett, M. L. *et al.* New tools for studying microglia in the mouse and human CNS. *Proceedings of the National Academy of Sciences* **113,** E1738–46 (2016).

53. Buttgereit, A. *et al.* Sall1 is a transcriptional regulator defining microglia identity and function. *Nat Immunol* **17,** 1397–1406 (2016).

54. Goldmann, T. *et al.* Origin, fate and dynamics of macrophages at central nervous system interfaces. *Nat Immunol* **17,** 797–805 (2016).

55. Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* **169,** 1276–1290.e17 (2017).

56. Haimon, Z. *et al.* Re-evaluating microglia expression profiles using RiboTag and cell isolation strategies. *Nat Immunol* **159,** 1312 (2018).

57. Bennett, F. C. *et al.* A Combination of Ontogeny and CNS Environment Establishes Microglial Identity. *Neuron* 1–23 (2018). doi:10.1016/j.neuron.2018.05.014

58. Cronk, J. C. *et al.* Peripherally derived macrophages can engraft the brain independent of irradiation and maintain an identity distinct from microglia. *J. Exp. Med.* **47,** jem.20180247–21 (2018).

59. Kierdorf, K. *et al.* Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat Neurosci* (2013). doi:10.1038/nn.3318

60. Song, W. M. & Colonna, M. The identity and function of microglia in neurodegeneration. *Nat Immunol* 1–11 (2018). doi:10.1038/s41590-018-0212-1

61. Mildner, A., Huang, H., Radke, J., Stenzel, W. & Priller, J. P2Y 12receptor is expressed on human microglia under physiological conditions throughout development and is sensitive to neuroinflammatory diseases. *Glia* **65,** 375–387 (2016).

62. Haynes, S. E. *et al.* The P2Y12 receptor regulates microglial activation by extracellular nucleotides. *Nat Neurosci* **9,** 1512–1519 (2006).

63. Takata, K. *et al.* Induced-Pluripotent-Stem-Cell-Derived Primitive Macrophages Provide a Platform for Modeling Tissue-Resident Macrophage Differentiation and Function. *Immunity* **47,** 183–198.e6 (2017).

64. Safaiyan, S. *et al.* Age-related myelin degradation burdens the clearance function of microglia during aging. *Nat Neurosci* **19,** 995–998 (2016).

65. ElAli, A. & Rivest, S. Microglia in Alzheimer's disease: A multifaceted relationship. *Brain, Behavior, and Immunity* **55,** 138–150 (2016).

66. Simard, A. R., Soulet, D., Gowing, G., Julien, J.-P. & Rivest, S. Bone Marrow-Derived Microglia Play a Critical Role in Restricting Senile Plaque Formation in Alzheimer's Disease. *Neuron* **49,** 489–502 (2006).

67. Krasemann, S. *et al.* The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity* **47,** 566–581.e9 (2017).

68. Holtman, I. R. *et al.* Induction of a common microglia gene expression signature by aging

and neurodegenerative conditions: a co-expression meta-analysis. *acta neuropathol commun* **3,** 31 (2015).

69. Varol, D. *et al.* Dicer Deficiency Differentially Impacts Microglia of the Developing and Adult Brain. *Immunity* **46,** 1030–1044.e8 (2017).

70. Yona, S. *et al.* Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* **38,** 79–91 (2013).

71. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell* **38,** 576–589 (2010).

72. Longair, M. H., Baker, D. A. & Armstrong, J. D. Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes. *Bioinformatics* **27,** 2453–2454 (2011).

73. Ferreira, T. A. *et al.* Neuronal morphometry directly from bitmap images. *Nat Meth* **11,** 982–984 (2014).

74. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol* **82,** 239–259 (1991).

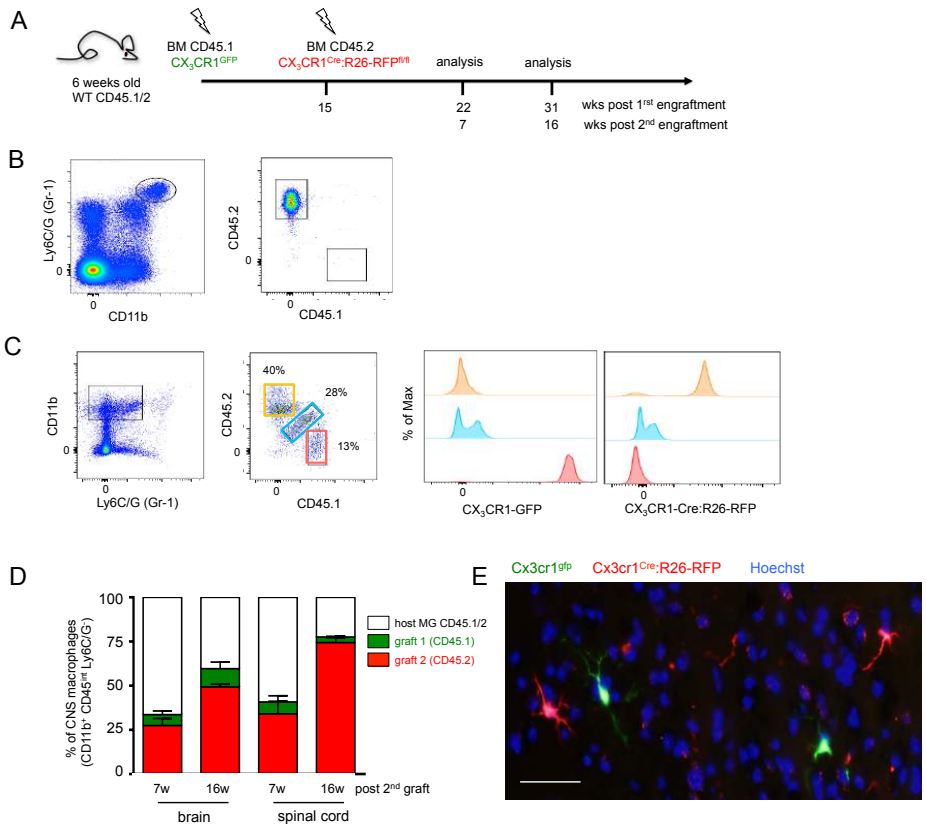


Fig. 1 BM-derived parenchymal brain macrophages accumulate over time post-irradiation and self-maintain independent of the circulation

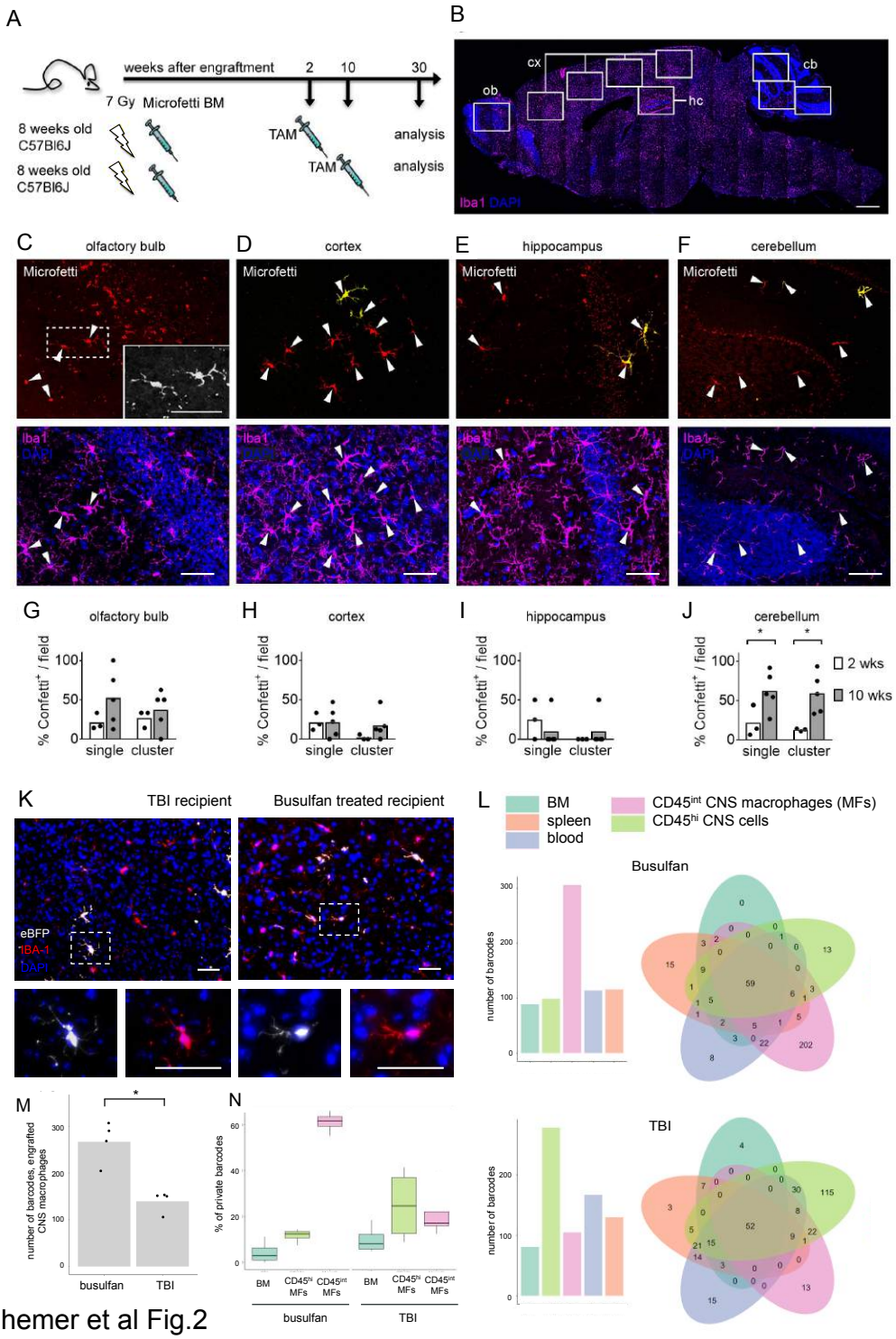
A) Schematic of tandem BM transfer protocol.

B) Flow cytometric blood monocyte analysis of chimera 16 weeks post-second transplantation.

C) Flow cytometric analysis of myeloid brain cells 16 weeks post second BMT revealing host microglia (CD45.1/2^{+}, blue), cells derived from the first graft (CD45.1^{+} GFP^{+}, red) and cells derived from the second graft (CD45.2^{+} RFP^{+}, orange).

D) Distribution of host and graft-derived cells out of the total Ly6C/G (Gr1)$^{-}$ CD45^{lo} CD11b^{+} cells in brain and spinal cord at two time points. Data are a summary of 6 mice.

E) Histological analysis 7 weeks post-second BMT revealing ramified GFP^{+} and RFP^{+} cells with microglia morphology (scale bar 20 μm). Representative picture.



Shemer et al Fig.2

Fig. 2 BM cells efficiently engraft recipient brains conditioned by irradiation or myelo-ablation and display clonal expansion

A) Fate mapping scheme for donor 'Microfetti' BM cells in lethally irradiated WT hosts. A single dose of TAM was applied at 2 or 10 weeks after BM reconstitution. BM chimeras were analyzed at 30 weeks after BM transplantation.

B) Representative sagittal brain section indicating the fields of view (rectangles) analyzed for IBA-1 (magenta) expressing graft microglia in the olfactory bulb (ob), cortex (cx), hippocampus (hc) and cerebellum (cb). DAPI (blue). Scale bar, 1 mm.

C-F) Representative images of Confetti$^+$ (yellow / red) Iba-1$^+$ (magenta) donor cells (arrowheads) in the (C) olfactory bulb (higher magnification in grey scale, inset), (D) cortex, (E) hippocampus, and (F) cerebellum. DAPI (blue). Scale bars, 50 µm (C-E) and 100 µm (F).

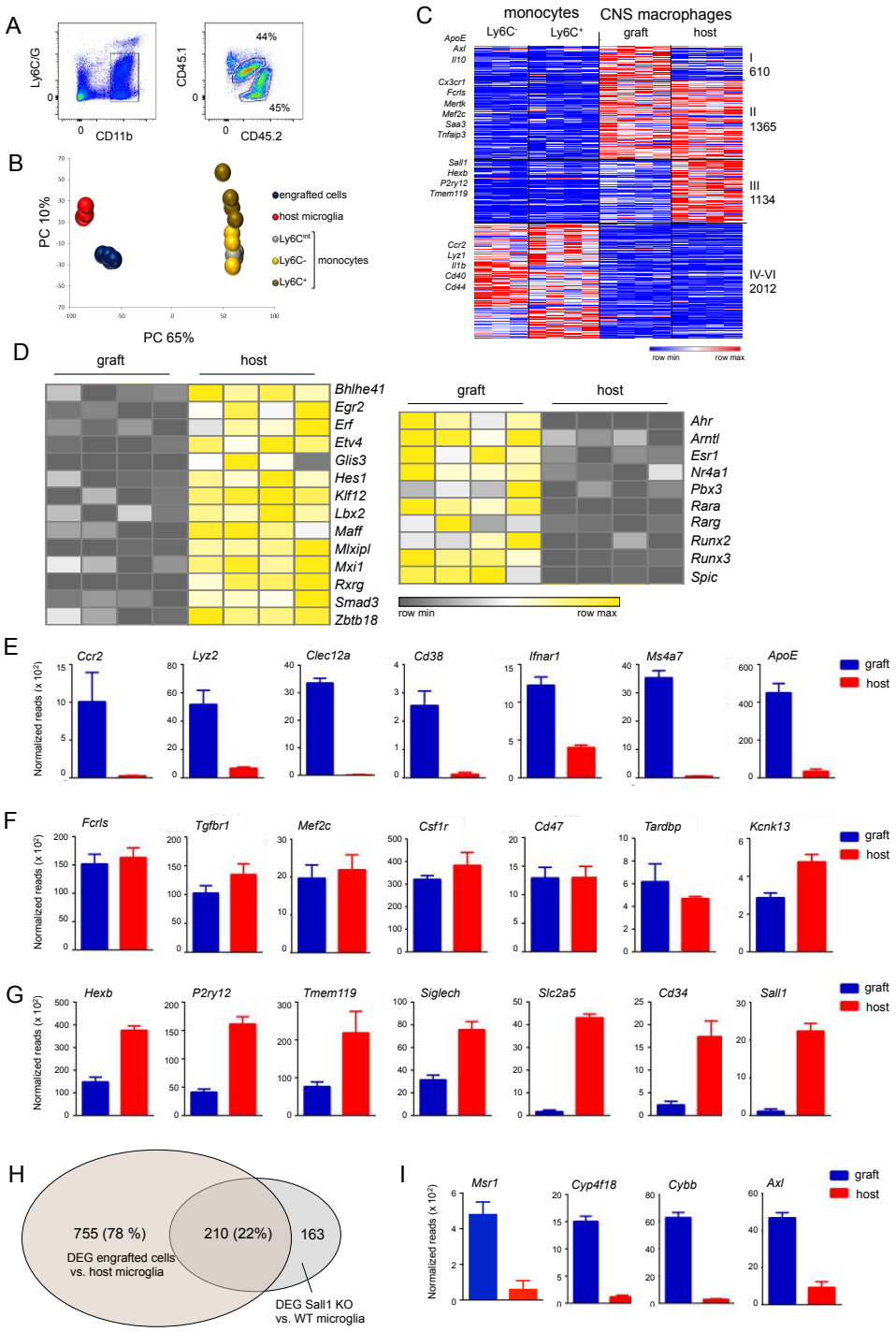
G-J) Frequency of Confetti$^+$ Iba-1$^+$ graft as single cells or clones in analyzed fields from TAM treatment groups of 2 weeks (white) and 10 weeks (gray) after BM transplantation. Each dot represents the mean quantification of one animal. At least four sections per animal were analyzed. Mann-Whitney test, * P = 0.0357, 0.0179, respectively in (J).

K) Representative images of cortical eBFP$^+$ (white) Iba-1$^+$ (red) graft-derived cells in brains of TBI and busulfan-conditioned mice. DAPI (blue). Higher magnification in insets. Scale bars, 50 µm.

L) Representative barcode analysis on DNA extracted from peripheral blood, bone marrow, spleen and sorted CD45^{int} brain macrophages and CD45^{high} cells of a busulfan- and a TBI-conditioned mouse. Barcode numbers for each sample are shown in the bar plot and the amount of shared barcodes per sample are displayed in the Venn diagrams.

M) Number of barcodes detected in engrafted macrophage samples of TBI- and busulfan conditioned animals. Each dot represents one animal (n=4 per group).

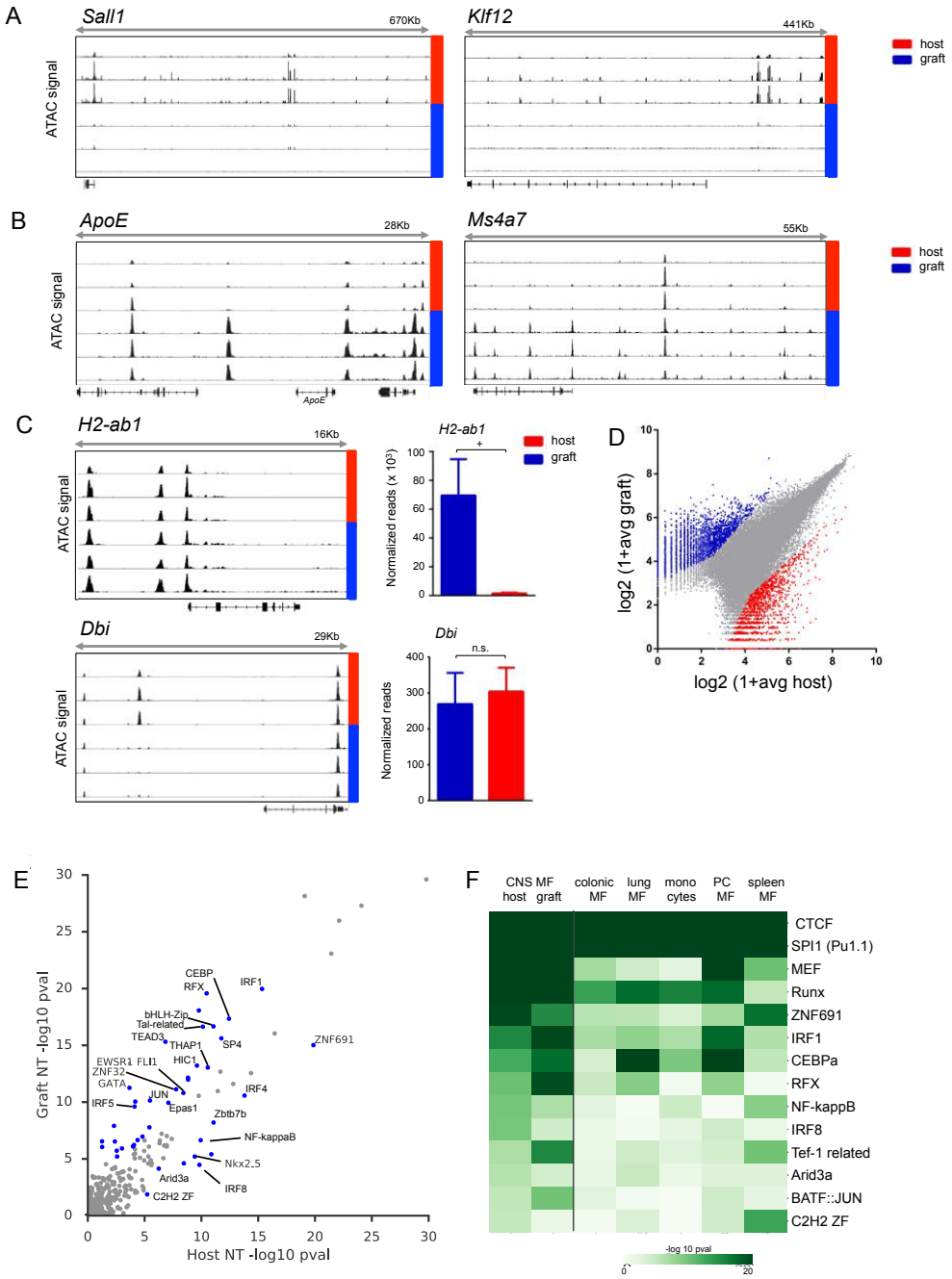
N) Bar graph showing percentages of barcodes private to grafted cells among CD45^{int} and CD45^{hi} CNS cells (i.e. non-parenchymal and parenchymal macrophages, respectively), as well as BM cells of TBI- and busulfan conditioned animals (n=4 per group).



Shemer et al Fig. 3

Fig. 3 Comparative transcriptome analysis of grafted cells and host microglia

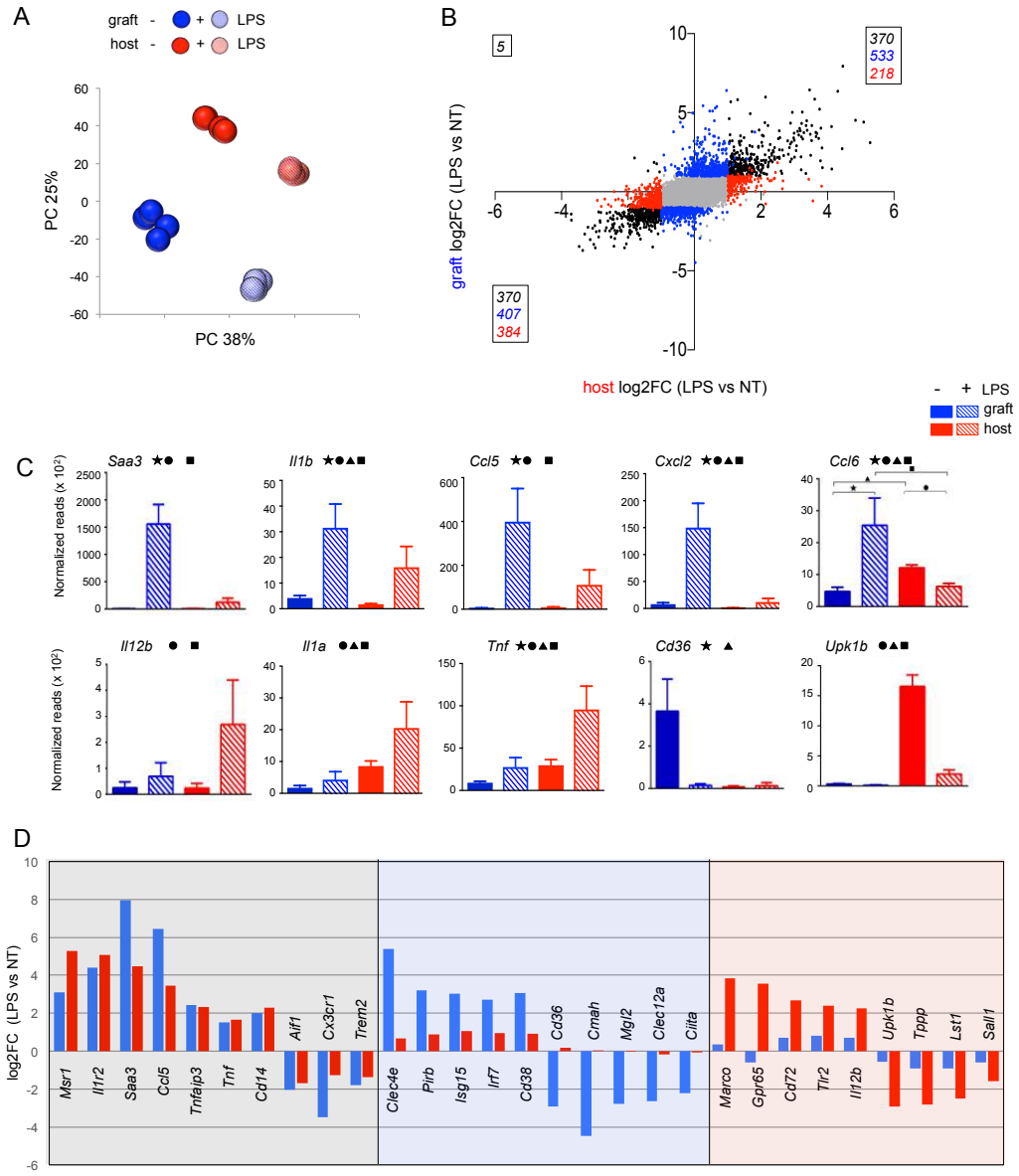
A) Gating strategy for isolation of CNS macrophages. Host microglia were defined as Ly6C/G$^-$ CD11b$^+$ CD45.2^{lo} cells; graft-derived cells were defined as Ly6C/G$^-$ CD11b$^+$CD45.1^{lo} cells.
B) Principal component analysis of transcriptomes of engrafted cells and host microglia and transcriptomes of monocytes subsets (Mildner et al. 2017).
C) Heat map of RNA seq data of engrafted cells and host microglia compared to transcriptomes of Ly6C$^-$ and Ly6C$^+$ monocyte subsets (Mildner et al. 2017). Analysis was restricted to genes, which showed a 2-fold difference and yielded p-value<0.05 between at least two populations.
D) Heatmap showing differential TF expression profiles of engrafted cells and host microglia.
E) Examples of significantly differential (log2FC>1 & p-value<0.05) gene expression enriched in engrafted cells. Graphs show normalized reads from RNA seq data of samples acquired in (A), n=4.
F) Examples of genes expressed in similar levels in host and engrafted cells. Graphs show normalized reads from RNA seq data of samples acquired in (A), n=4. None of the genes were differentially expressed (log2FC<1) and *Tgfbr1* and *Kcnk13* were of low significance (p-value<0.05) that did not meet our threshold.
G) Examples of significantly differential (log2FC<-1, p-value<0.05) gene expression enriched in host microglia. Graphs show normalized reads from RNA seq data of samples acquired in (A), n=4.
H) Venn diagram showing overlap of genes differentially expressed by WT and Sall1-deficient microglia (Buttgereit et al., 2016), and genes differentially expressed by host and engrafted brain macrophages.
I) Examples of gene expression of engrafted and host cells of genes expressed in non-CNS macrophages or Sall1-deficient microglia.



Shemer et al Fig. 4

Fig. 4 Comparative epigenome analysis of grafted cells and host microglia

A) IGV tracks of *Sall1* and *Klf2* loci showing ATAC signals in host (red) but not engrafted (blue) cells. N=3
B) IGV tracks of *ApoE* and *Ms4a7* loci showing ATAC signals in engrafted cells (blue) but not host (red) microglia. N=3
C) ATAC seq IGV tracks (left, n=3) and normalized RNA seq reads (right, n=4) of *H2-ab1* and *Dbi* in host microglia (red) and engrafted cells (blue). + – p-value<10^{-5}, n.s. – p-value>0.05.
D) Analysis of all 58,947 detected ATAC peaks, from which 1,506 peaks and 2,176 peaks displayed >4 fold significant (p-value<0.01) enrichment in host microglia and engrafted cells, respectively.
E) Comparison of motif significance in resting host and grafted cells. P-values were calculated using TBA models trained on intergenic peaks from host and grafted cells. Significant motifs that show a large difference (p<10e-5, log likelihood ratio >= 2) are indicated in blue points.
F) Heatmap of the significance of motifs in various myeloid cell types and host cells. The intensity of the color indicates greater significance (-log 10 p-value) of each motif.



Shemer et al Fig. 5

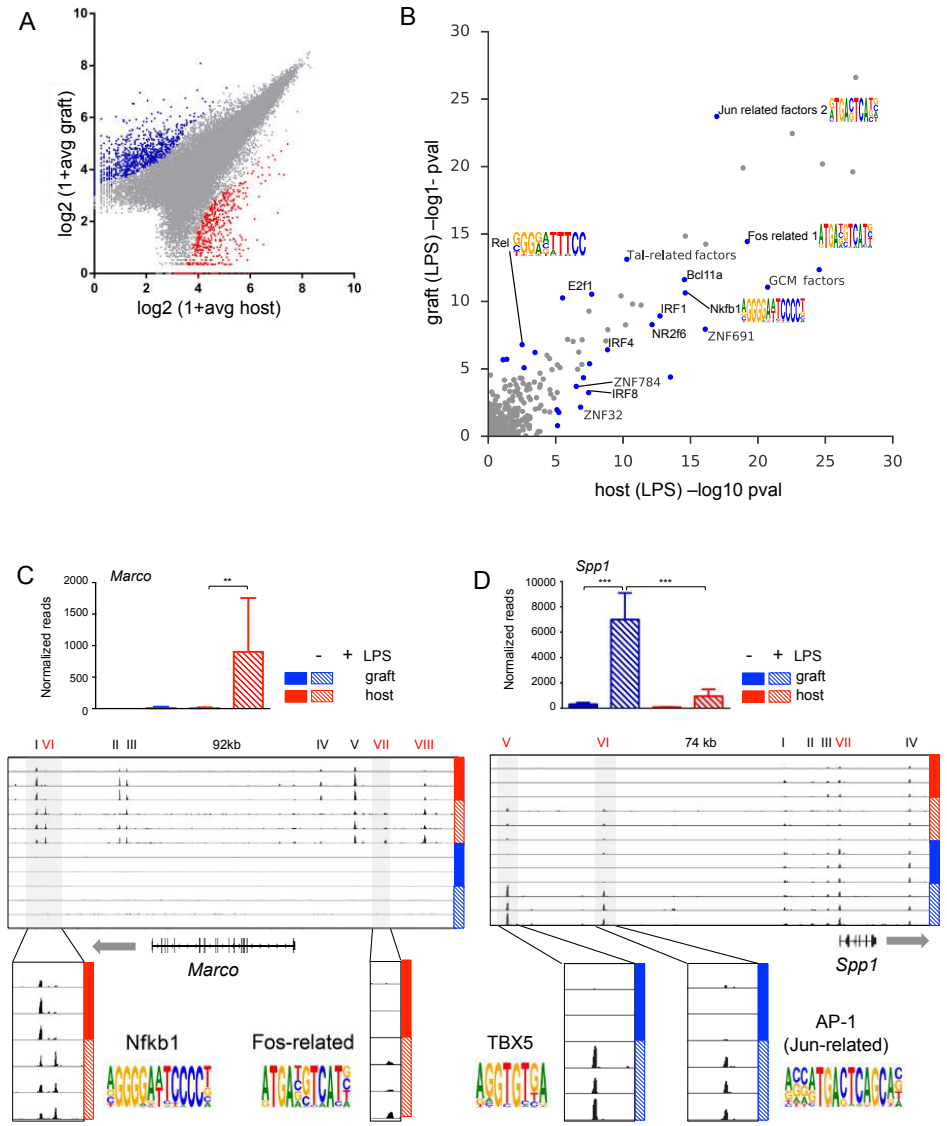
Fig.

5 Distinct LPS responses of engrafted cells and host microglia

A) PC analysis of RNA seq data of graft and host microglia in steady state and 12h post-LPS.
B) Expression analysis of grafted cells and host microglia in steady state and 12h post-LPS by RNA seq.
C) Examples of genes expression of graft and host microglia in steady state and 12h post-LPS.

   Significance is indicated by the symbols.

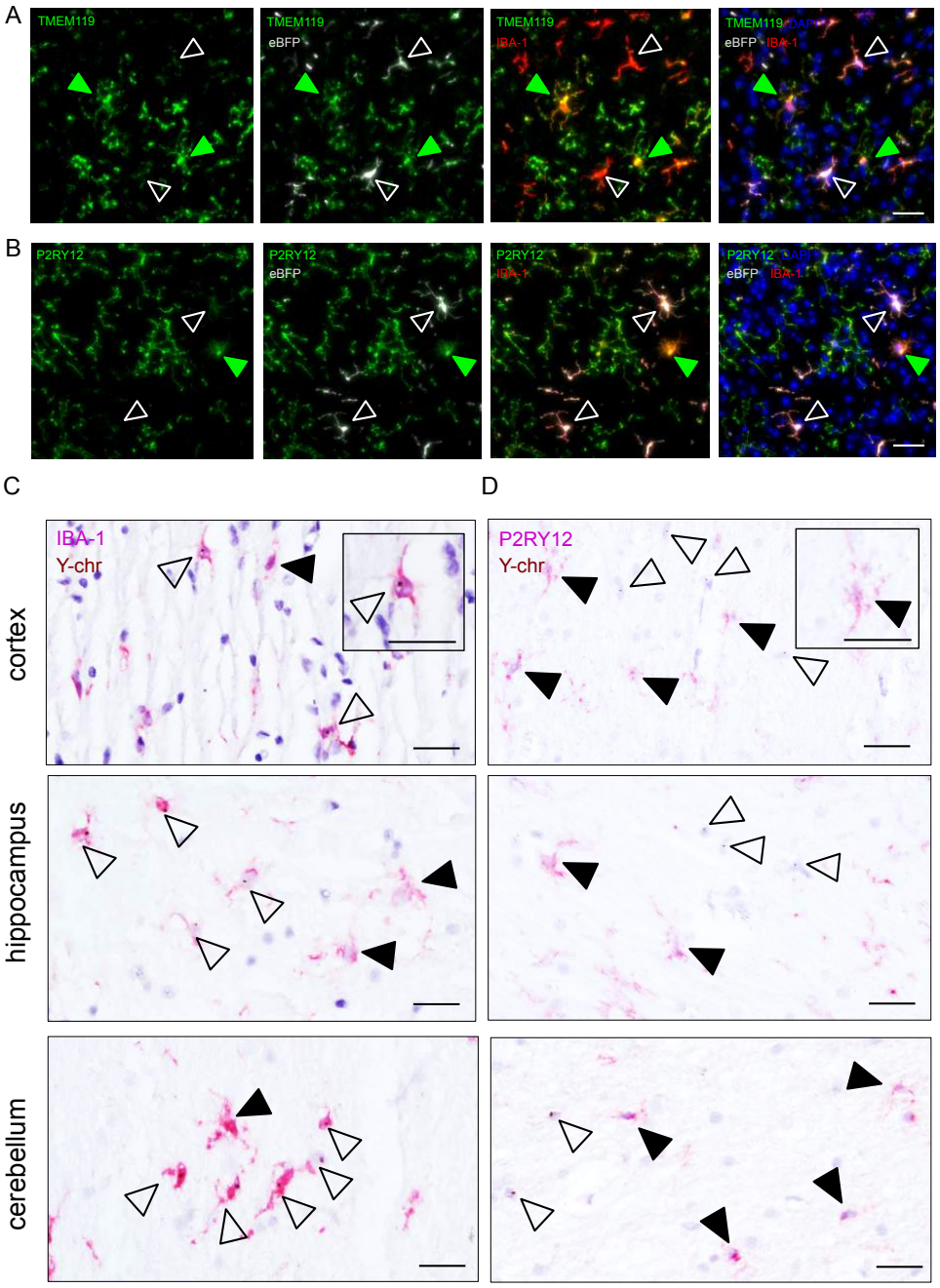
D) Fold expression change of selected genes significantly different (absolute value of log2FC > 1, p-value<0.05) following challenge of grafted cells (middle) or of host microglia (right), as well as genes displaying comparable up- and down-regulation in both engrafted cells and host microglia (left).



Shemer et al Fig. 6

Fig. 6. Comparative epigenome analysis of grafted cells and host microglia post-LPS

A) Analysis of all 46,485 detected ATAC peaks, from which 552 peaks and 841 peaks displayed >4 fold significant (p-value<0.01) enrichment in host microglia and engrafted cells isolated from challenged mice, respectively.
B) Comparison of motif significance in activated/challenged host and grafted cells. P-values were calculated using TBA models trained on intergenic ATAC-seq peaks from host and grafted cells. Significant motifs that show a large difference (p<10e-5, log likelihood ratio >= 2) are indicated in blue points. Motif logos visualizing the position frequency matrix of NF-kappaB and AP-1 motifs are annotated.
C) Challenge induced alterations in *Marco* locus. Normalized sequence reads of *Marco* mRNA in engrafted cells and host microglia isolated from LPS challenged and unchallenged BM chimeras (left); Normalized ATACseq profiles of *Marco* locus (right), with enlarged areas highlighting induced ATACseq peaks and predicted motifs.
D) Challenge induced alterations in *Spp1* locus. Normalized sequence reads of *Spp1* mRNA in engrafted cells and host microglia isolated from LPS challenged and unchallenged BM chimeras (left); Normalized ATACseq profiles of *Spp1* locus (right), with enlarged areas highlighting induced ATACseq peaks and predicted motifs.

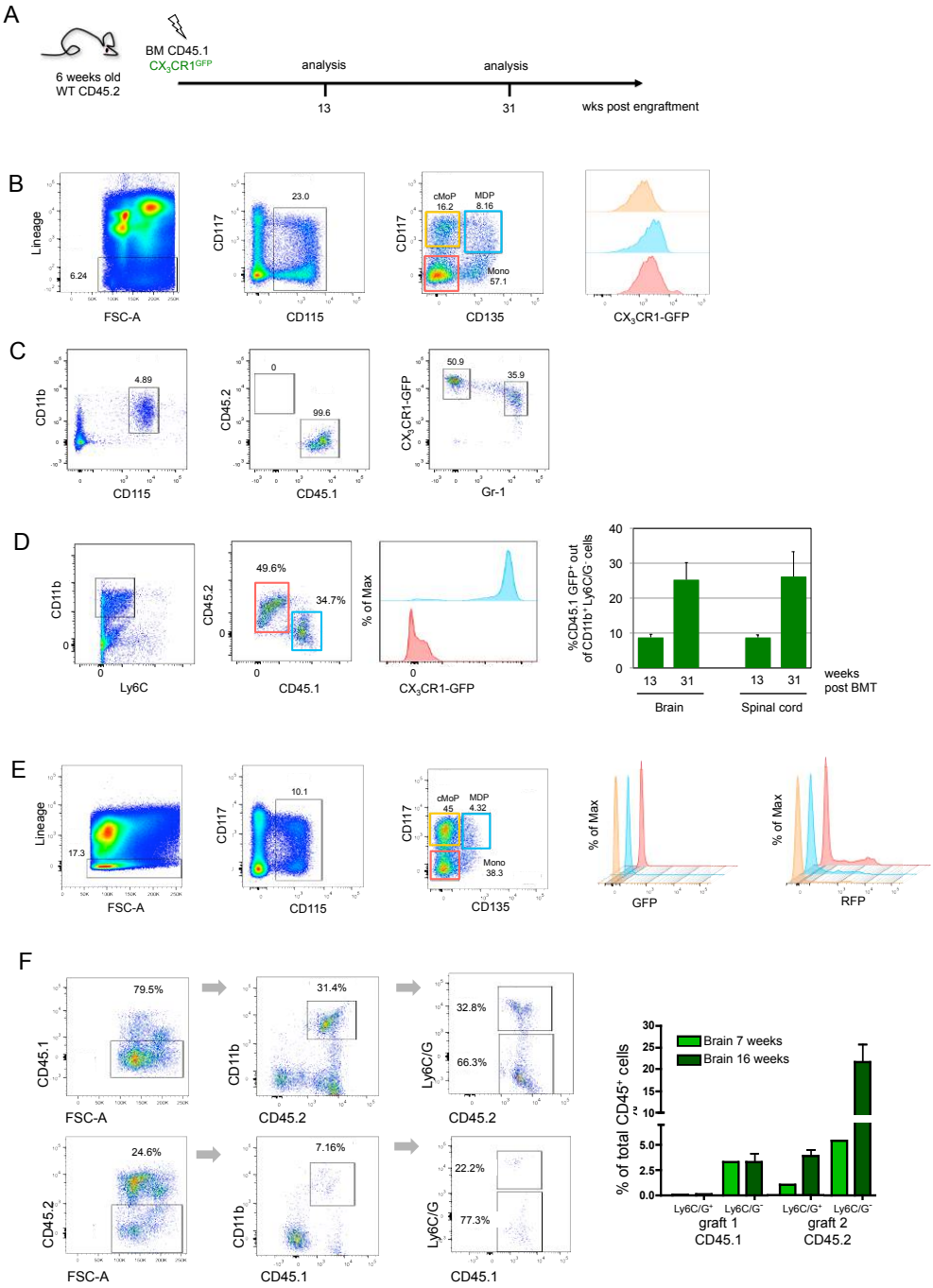


Shemer et al Fig. 7

Fig. 7. Comparative protein expression analysis of grafted cells and host microglia in mouse and human chimeras.

A-B) Expression of host microglia-specific markers TMEM119 (A) and P2RY12 (B) in the cortex of mice that received lineage negative BM carrying the lentiviral construct conferring eBFP expression. Host microglia (green) and donor cells (white) are indicated with arrowheads. Iba-1 immuno-histochemistry for microglia (red). DAPI nuclear counterstain (blue). Scale bars, 30 μm.

C-D) Representative images of cortical, hippocampal and cerebellar sections from female patients that received male donor BM grafts carrying the Y-chromosome (Y-chr). Host microglia are indicated by solid arrow heads and donor cells are shown by open arrowheads. Immunohistochemistry (red) of IBA-1 (C) and P2RY12 (D) combined with *in situ* hybridization of Y-chr (brown). Insets show a single grafted cell at higher magnification. Scale bars, 30 μm.



Supplementary Fig. 1

Supplementary Fig. 1 BM-derived parenchymal brain macrophages accumulate over time post-irradiation

A) Schematic of BM transfer protocol.

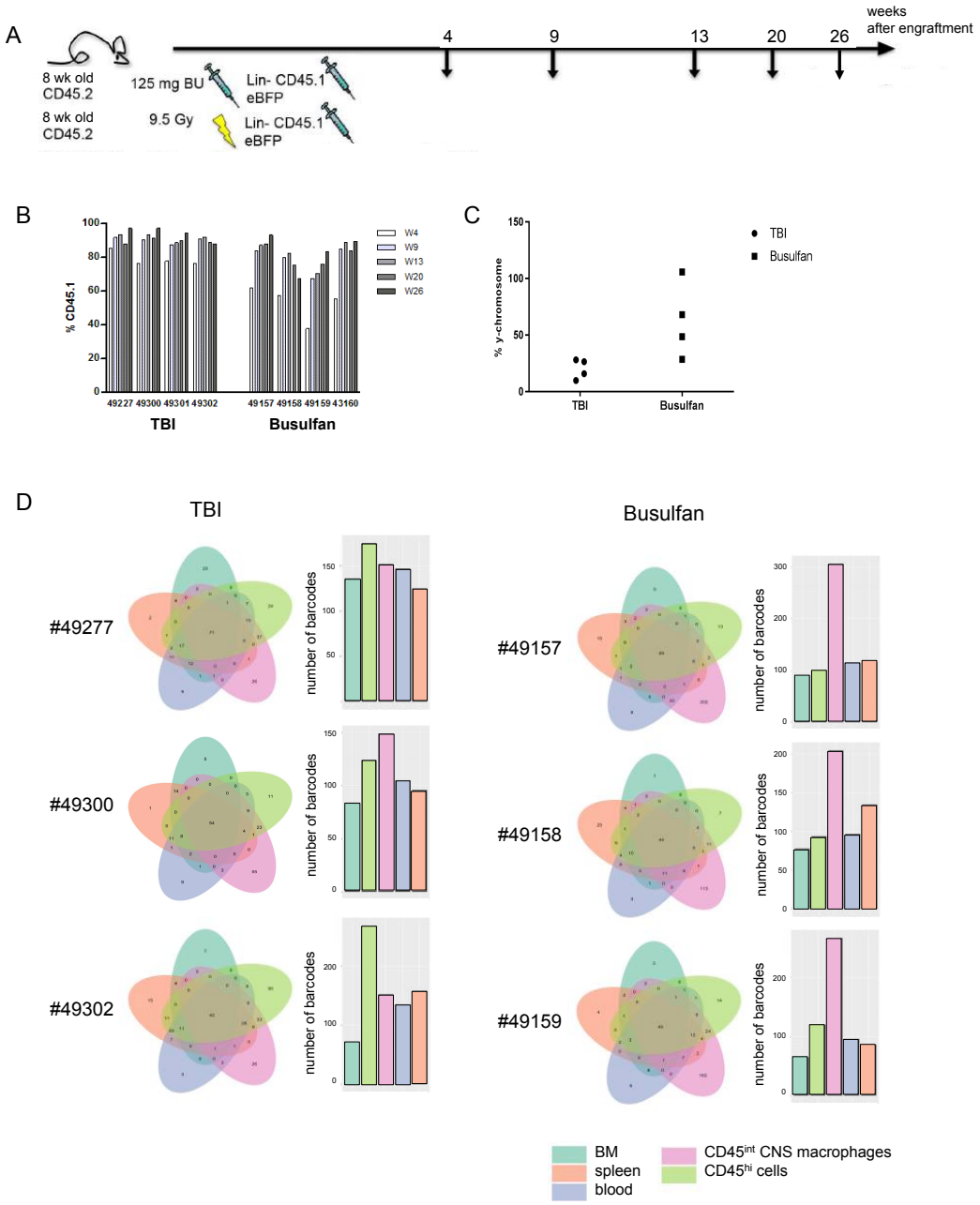
B) Flow cytometry analysis of myeloid BM progenitor compartment of chimeras 13 weeks post-BMT.

C) Flow cytometry analysis of blood monocytes of chimeras 13 weeks post-BMT.

D) Flow cytometry analysis of brain cells of chimeras 13 weeks post-BMT. Of note, $Ly6C^{-}CD11b^{+}$ cells comprise host microglia (CD45.2) and BM-derived cells (GFP^{+} CD45.1).

E) Analysis of myeloid BM progenitor compartment of chimeras 7 weeks post second BMT. Note that none of the myeloid progenitors expressed GFP (GFP histogram), while MDP, cMOP and monocytes do show RFP signal (RFP histogram).

F) Flow cytometric analysis of brain cells 16 weeks post second BMT for distribution of first and second graft-derived myeloid cells ($CD11b^{+}$ Ly6C/G $(Gr1)^{+}$ and $CD11b^{+}$ Ly6C/G $(Gr1)^{-}$). Bar graph shows distribution of cells derived from first and second graft out of the total $CD45^{+}$ cells in brains at two time points. Data are a summary of 3 mice per time point.



Supplementary Fig. 2

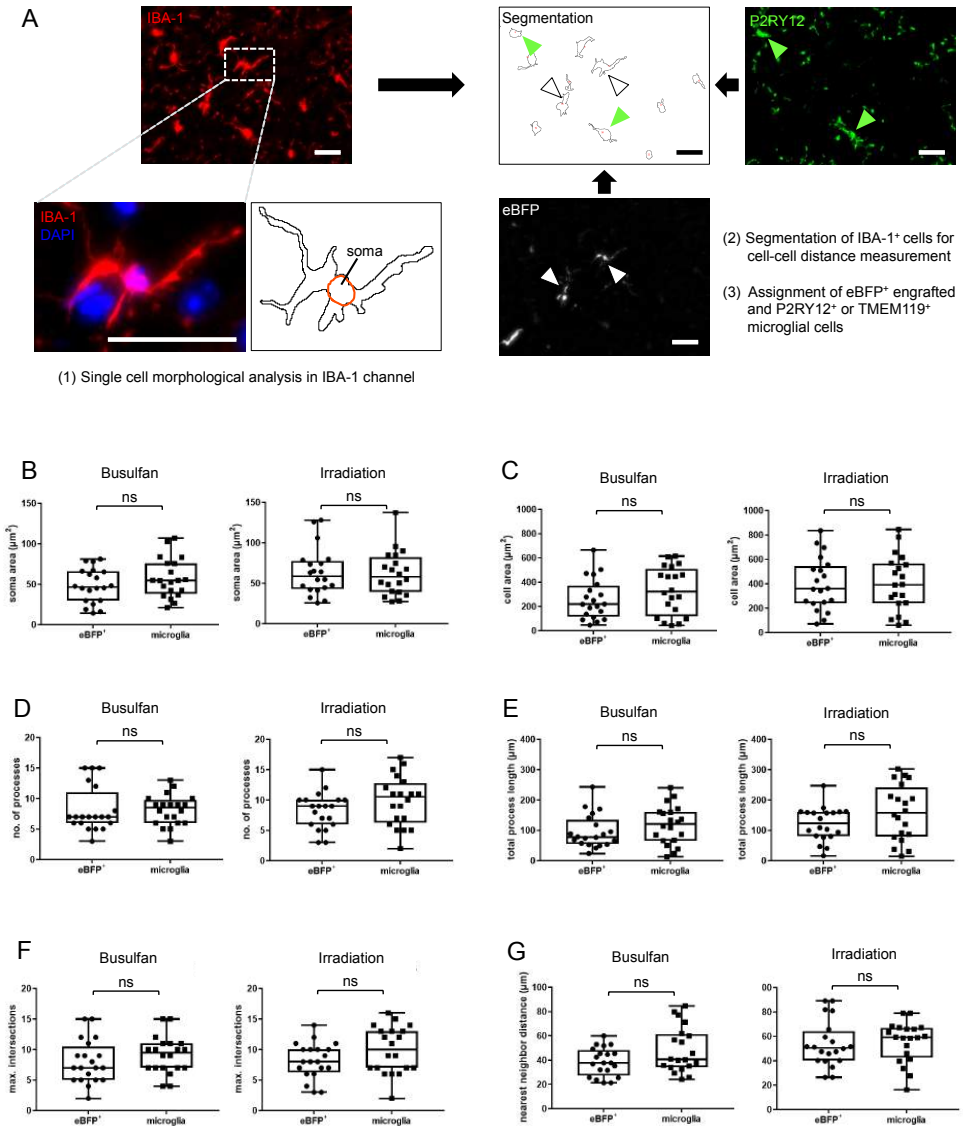
Supplementary Fig. 2 Clonal analyses on TBI and Busulfan treated animals.

A) Experimental set-up: Lineage-negative cells from 8-weeks old male donors were transduced with lentiviral vectors carrying eBFP as a fluorescent reporter as well as BC32-eBFP barcode system prior to transplantation into 8-weeks old female recipients conditioned either with total body irradiation (TBI) or chemotherapy (Busulfan). Peripheral blood samples were taken every 4-6 weeks post transplantation to monitor engraftment. 6 months post transplantation, the animals were sacrificed and peripheral blood, BM, spleen and the brain were taken for clonal analyses.

B) Flow cytometry on peripheral blood samples to measure the chimerism (CD45.1 on graft cells).

C) Chimerism of sorted microglia cells as detected by Y-chromosome-specific digital droplet PCR.

D) Barcode analyses on DNA extracted from peripheral blood, bone marrow, spleen and sorted $CD45^{int}$ brain macrophages and $CD45^{high}$ CNS cells. Numbers of barcodes for the samples are shown in the bar plots and shared barcodes of the respective sample are displayed in the Venn diagrams.

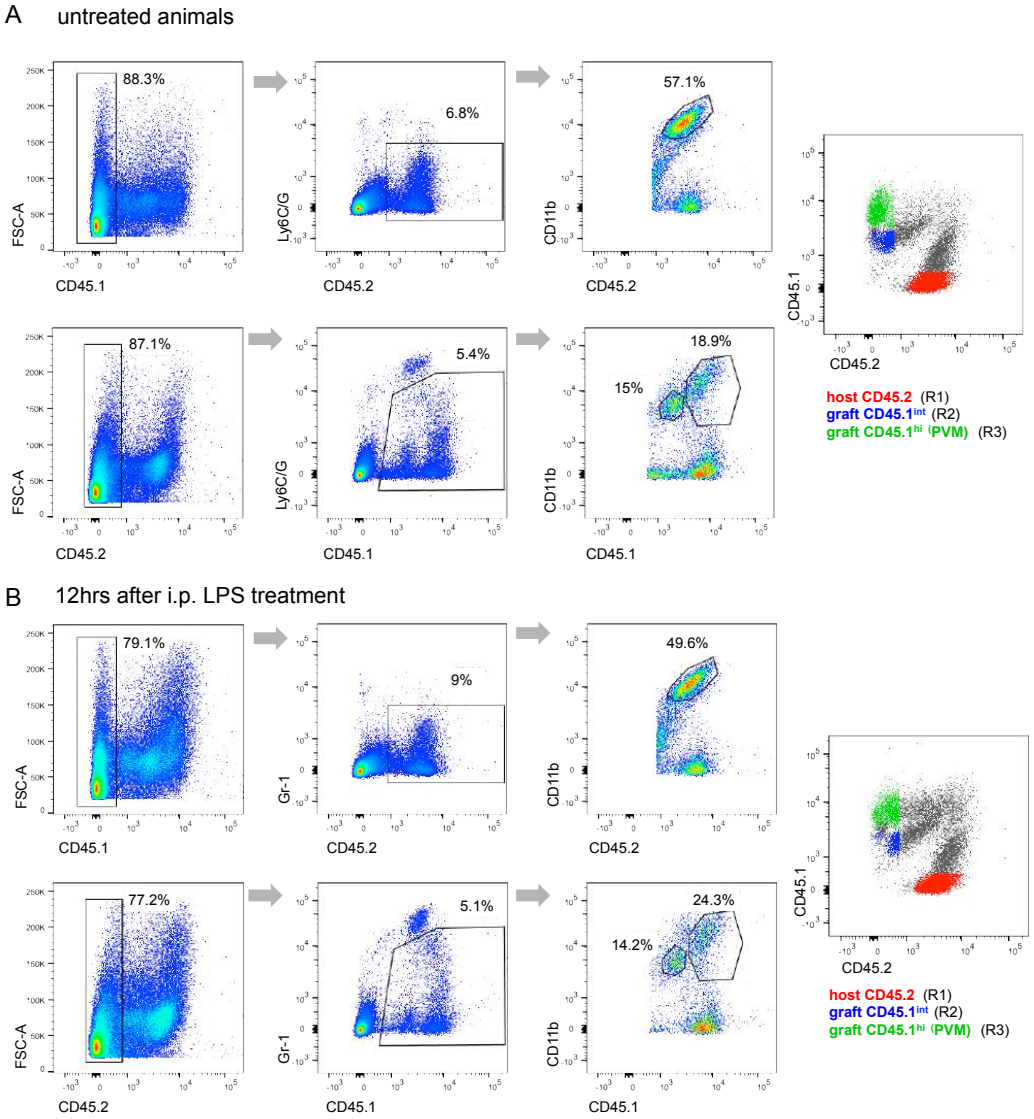


Supplementary Fig. 3

Supplementary Fig. 3 Morphological and intercellular distance measurements of host and donor cells from recipients of lineage negative BM that carried the lentiviral construct conferring eBFP expression.

A) Workflow of measurements shows that all cortical microglial cells were segmented and randomly sampled for single cell morphological analysis in the Iba-1 (red) channel. DAPI nuclear counterstain (blue) was used to indicate soma location. For unblinding of cell type, immunolabel for P2RY12 (green; filled arrowheads) or TMEM119 (not shown) identified host microglial cells, while eBFP expression (white; open arrowheads) indicated donor cells. Scale bars, 30 μm.

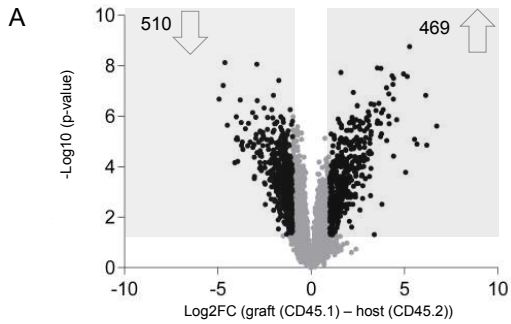
B-G) In maximum intensity projection images, single cells were analyzed for (B) area of cell soma, (C) area covered by soma and processes, (D) number of processes, (E) total length of the processes, (F) maximum number of intersections of the processes determined by Sholl analysis, and (G) nearest neighbor intercellular distance. All data sets are normally distributed (D'Agostino-Pearson test, $p > 0.05$), except the total process length of $eBFP^+$ donor cells in Busulfan-conditioned mice (D'Agostino-Pearson test, $p = 0.0358$). For each brain conditioning paradigm (Busulfan or total body irradiation), the parameters of 20 cells per cell type were compared by two-tailed unpaired *t*-test. ns, not significant.

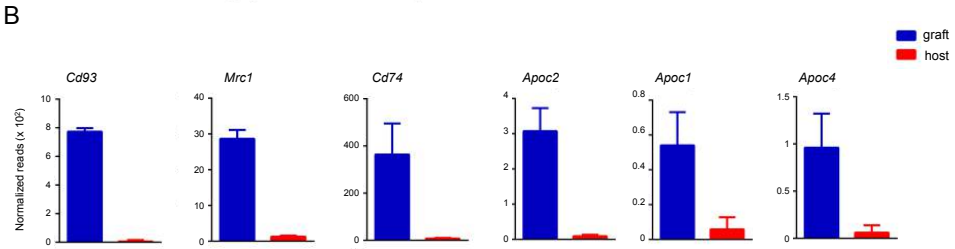


Supplementary Fig. 4 Flow cytometry analysis of chimeras used for isolation of cells for transcriptome epigenome analysis

A) Representative flow cytometry analysis of brain of [CD45.1 > CD45.2] chimera, 9 months after engraftment. Note presence of graft derived- parenchymal macrophages ($CD45.1^{int}$) and perivascular macrophages ($CD45.1^{hi}$), but absence of the latter from the $CD45.1^{+}$ host compartment. Dot blot on the right indicates respective population in the sort gate (Fig. 3a)

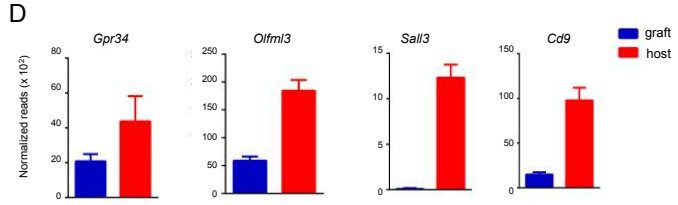
B) Representative flow cytometry analysis of brain of [CD45.1 > CD45.2] chimera, 9 months after engraftment and 12 hours after intra-peritoneal LPS treatment. Dot blot on the right indicates respective population in the sort gate (Fig. 3a).





C

	host 1	host 2	host 3	host 4	host 5	graft 1	graft 2	graft 3	graft 4	graft 5	non-irradiated MG 1	non-irradiated MG 2	non-irradiated MG 3
host 1	1												
host 2	0.988472	1											
host 3	0.974321	0.993990647	1										
host 4	0.986449	0.997411836	0.994455	1									
host 5	0.959172	0.986139	0.989819	0.988375	1								
graft 1	0.431397	0.446284	0.450021	0.433159	0.420734	1							
graft 2	0.20391	0.220991	0.233341	0.211148	0.204299	0.9501446	1						
graft 3	0.533481	0.542920	0.539857	0.530383	0.512875	0.961605405	0.857912195	1					
graft 4	0.374389	0.389396	0.395167	0.377384	0.365958	0.95722337	0.902023319	0.922959506	1				
graft 5	0.331718	0.346608	0.352759	0.336069929	0.326516414	0.97286904	0.939637803	0.928676741	0.986788168	1			
non-irradiated MG 1	0.958512597	0.979276061	0.976153859	0.977068043	0.976856912	0.52408286	0.311645525	0.603922276	0.459988027	0.429121818	1		
non-irradiated MG 2	0.95303496	0.975811799	0.971227357	0.973408692	0.978910242	0.512409546	0.299483711	0.596356184	0.44822863	0.41930029	0.99712232	1	
non-irradiated MG 3	0.960912618	0.98000665	0.977206527	0.97853232	0.975458965	0.530242354	0.318278007	0.611133817	0.465887866	0.433369758	0.99776659	0.995292889	1



Supplementary Fig. 5. Comparison of transcriptomes of engrafted cells and host microglia

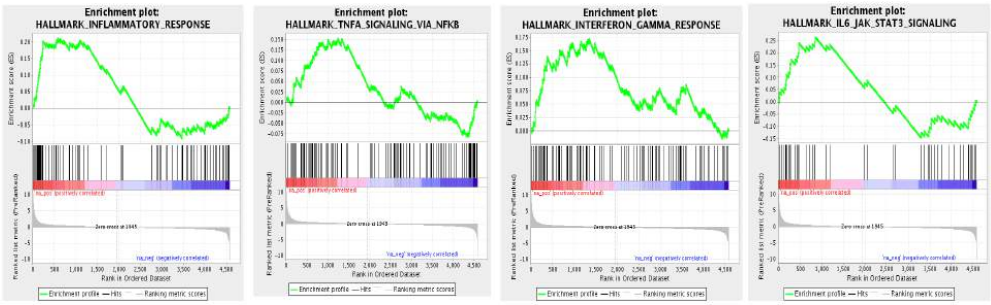
A) Volcano plot of RNA seq data of samples acquired in Fig. 3A, showing up- and down regulated genes in the graft-derived cells relative to host microglia. Analysis was restricted to expressed genes. Dark dots indicate genes which showed a 2-fold difference and yielded p-value<0.05 between the two populations, N=4.

B) Correlation matrix of transcriptomes of engrafted cells and host microglia with transcriptome of microglia isolated from brains of age-matched non-irradiated C57Bl6 WT mice.

C) Examples of gene expression enriched in engrafted cells. Graphs show normalized reads from RNA seq data of samples acquired in Fig. 3A, n=4.

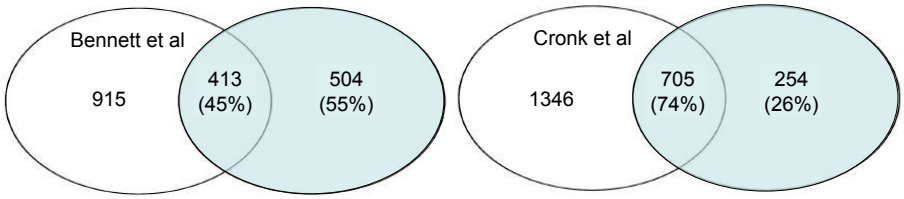
D) Examples of gene expression enriched in host microglia. Graphs show normalized reads from RNA seq data of samples acquired in Fig. 3A, n=4.

A



FDR q –value 0.005230476 FDR q –value 0.0031200416 FDR q –value 0.00117948055 FDR q –value 0.009422488

B

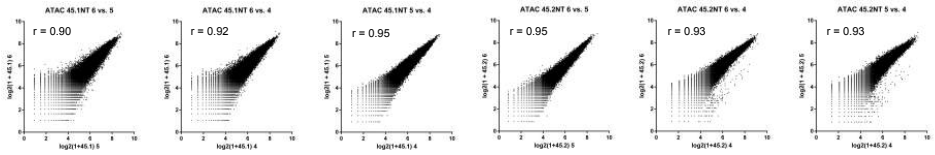


Supplementary Fig. 6. Pathway analysis and transcriptome comparisons

A) GSEA analysis of transcriptomes of HSC-derived brain macrophages and microglia.

B) Comparative analysis of lists of genes differentially expressed by BM graft derived cells and host microglia in this study to the data retrieved from the study by Kipnis and colleagues [58] and Bennett et al. [57]. Blue area represents DEG of engrafted macrophages and host microglia of this study.

A



B

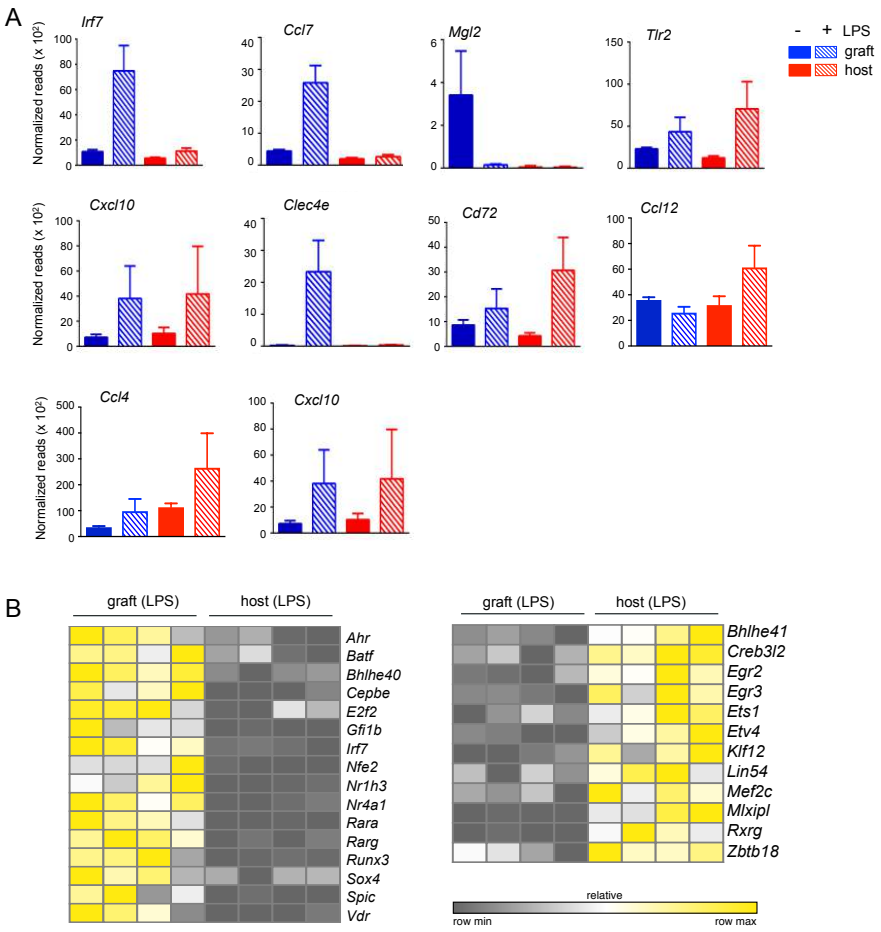
-log10 (p-value)

motif	graft	host
BCL11A	40.04648	23.06527
CEBPA	32.43423	21.85961
CTCF	50	50
CTCFL	50	41.91764
ETV6	50	50
Ets-related_factors_1_merged	43.98128	33.19889
Ets-related_factors_3_merged	50	50
Regulators_of_differentiation_1_merged	50	50
Runt-related_factors_2_merged	38.75022	24.48596
ZEB1	50	49.99995

Supplementary Fig. 7. ATACseq data of engrafted cells and host microglia isolated from brains of unchallenged BM chimeras.

A) Correlation analysis of ATACseq samples prepared from sorted grafted cells and host microglia.

B) List of highly abundant motifs identified in sequences coinciding with ATACseq peaks (-log10 (p-value) > 66.37).



Supplementary Fig. 8. Comparison of engrafted cells and host microglia isolated from brains of BM chimeras subjected to a peripheral LPS challenge

A) Examples of genes expression of grafted cells and host microglia in steady state and 12h post-LPS.

B) Heatmap showing differential TF expression profiles of grafted cells and host microglia isolated from LPS challenged chimeras.

A

Upstream Regulator	Expr Log Ratio (graft / host)	Predicted Activation State	Activation z-score	p-value of overlap	Mechanistic Network
TGFB1	-0.464		-0.029	1.58E-28	276 (18)
LPS			1.936	2.40E-28	299 (17)
IL4			0.083	4.74E-24	254 (15)
IFNG			1.713	1.08E-22	241 (19)
IL10RA	-0.334		-1.72	7.44E-22	193 (19)
TNF	-1.59		0.107	5.19E-21	264 (16)
dexamethasone			0.274	4.86E-20	285 (19)
IL6	-2.33		1.384	2.36E-19	261 (17)
APP	0.337		0.192	2.51E-19	244 (17)
CSF3	0		0.791	3.61E-17	234 (22)
CSF2			1.015	4.18E-17	283 (19)
IL10	5.02		-1.358	6.89E-17	237 (19)
poly rI:rC-RNA		Activated	2.711	9.02E-17	235 (18)
IL1B	1.37		1.803	1.15E-16	247 (16)
IL2		Activated	2.071	3.18E-16	251 (18)

B

Upstream Regulator	Expr Log Ratio (graft / host)	Predicted Activation State	Activation z-score	p-value of overlap	Mechanistic Network
LPS		Activated	6.292	2.38E-64	324 (15)
IFNG		Activated	5.652	4.75E-53	289 (14)
TNF	-1.71	Activated	5.042	2.62E-43	277 (14)
IL1B	1.03	Activated	4.395	1.70E-41	322 (14)
TGFB1	0.332		0.544	3.00E-37	311 (19)
IL4		Activated	2.937	6.06E-33	309 (15)
dexamethasone			1.166	1.06E-32	321 (16)
IL10RA	-0.159	Inhibited	-4.971	8.40E-31	217 (13)
TNFSF11		Activated	2.416	1.43E-29	259 (14)
IL13			1.037	1.87E-29	267 (15)
STAT1	0.536	Activated	4.556	5.92E-29	262 (16)
STAT3	0.14		1.69	9.19E-29	275 (18)
IL6	-0.955	Activated	3.462	2.84E-28	296 (16)
IRF3	-0.219	Activated	4.755	1.28E-27	206 (13)
E. coli B4 LPS		Activated	3.406	4.25E-27	243 (17)

C

Upstream Regulator	Expr Log Ratio (graft NT / LPS)	Predicted Activation State	Activation z-score	p-value of overlap	Mechanistic Network
LPS		Activated	10.409	2.79E-76	515 (13)
IFNG		Activated	7.137	5.98E-63	447 (13)
IL1B	2.85	Activated	7.232	1.47E-55	455 (13)
TNF	1.54	Activated	8.266	3.94E-51	461 (13)
IL4		Activated	3.135	5.43E-43	527 (17)
TGFB1	-0.71	Activated	3.15	1.20E-42	510 (18)
dexamethasone			0.194	1.84E-38	602 (17)
TNFSF11		Activated	4.885	1.67E-34	462 (13)
IL10RA	-0.585	Inhibited	-6.758	1.64E-30	344 (16)
poly rI:rC-RNA		Activated	6.779	2.74E-29	461 (16)
beta-estradiol		Activated	2.073	2.10E-28	649 (20)
TP53	-0.9		1.761	3.08E-28	489 (19)
CD40LG			1.886	7.21E-28	409 (15)
IL6	2.54	Activated	3.755	1.02E-27	458 (13)
PD98059		Inhibited	-3.87	2.02E-27	550 (19)

D

Upstream Regulator	Expr Log Ratio (host NT/ LPS)	Predicted Activation State	Activation z-score	p-value of overlap	Mechanistic Network
LPS		Activated	7.457	5.68E-33	367 (14)
IL1B	3.19	Activated	5.235	1.20E-26	315 (13)
IFNG		Activated	5.276	2.05E-25	317 (14)
TNFSF11		Activated	3.984	2.88E-23	290 (11)
poly rI:rC-RNA		Activated	5.368	1.04E-22	323 (15)
TNF	1.66	Activated	5.972	1.65E-21	310 (13)
TLR4	-1.21	Activated	4.144	1.29E-19	283 (14)
TGFB1	-1.51	Activated	2.604	3.02E-19	374 (16)
dexamethasone			0.666	5.94E-19	385 (17)
E. coli B5 LPS		Activated	4.473	1.72E-18	308 (13)
IL4			-0.634	6.98E-18	325 (17)
S. minnesota R595 LPSs		Activated	4.95	7.57E-18	285 (14)
TICAM1	-0.095	Activated	5.04	1.02E-16	227 (16)
methylprednisolone		Activated	2.942	1.54E-16	388 (19)
APP	1.26	Activated	4.786	2.25E-16	371 (16)

Supplementary Fig. 9. Ingenuity Pathway analyses (IPA)

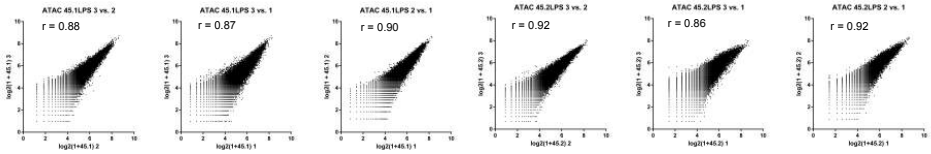
A) IPA analysis of genes differentially expressed by engrafted macrophages and host macrophages isolated from untreated BM chimeras.

B) IPA of transcriptomes of HSC-derived engrafted parenchymal brain macrophages and host microglia isolated from animals 12hrs after peripheral LPS challenge.

C) IPA of genes differentially expressed by engrafted macrophages isolated from untreated BM chimeras or animals 12hrs after peripheral LPS challenge.

D) IPA of genes differentially expressed by host microglia isolated from untreated BM chimeras or animals 12hrs after peripheral LPS challenge.

A



B

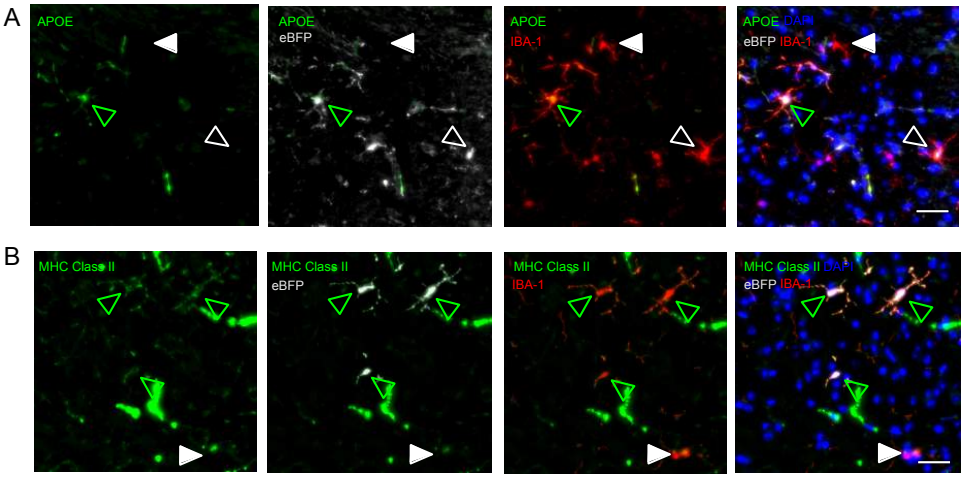
-log10 (p-value)

	graft	host
CEBPA	49.329186	49.9673163
CTCF	50	50
ETV6	23.1960544	35.2170259
Ets-related_factors_3_merged	50	50
ZEB1	41.0636486	39.8957154

Supplementary Fig. 10. ATACseq data of engrafted cells and host microglia isolated from mice challenged with LPS

A) Correlation analysis of ATACseq samples prepared from sorted grafted cells and host microglia isolated from LPS challenged BM chimeras.

B) List of highly abundant motifs identified in sequences coinciding with ATACseq peaks (-log10 (p-value) > 50.86) in grafted cells and microglia isolated from LPS challenged BM chimeras.



Supplementary Fig. 11. Comparative protein expression analysis of grafted cells and host microglia in mouse chimeras.

Expression of the graft-specific markers APOE (A) and MHC class II (B) (green) in the cortex of mice that received lineage negative BM carrying the lentiviral construct conferring eBFP expression. Host (filled) and donor (open) cells are indicated with respective arrowheads. Iba-1 immuno-histochemistry for microglia (red). DAPI nuclear counterstain (blue). Scale bars, 30 µm.