

Polygenic risk for Alzheimer's disease shapes microglial inflammatory and antigen-presentation programs in vivo

Bart De Strooper

b.strooper@ucl.ac.uk

VIB/KU Leuven Center for Brain & Disease Research and UKDRI/University College London, UK
<https://orcid.org/0000-0001-5455-5819>

Annerieke Sierksma

VIB Center for the Biology of Disease, Leuven, Belgium Center for Human Genetics, Universitaire ziekenhuizen and LIND, KU, Leuven, Belgium. <https://orcid.org/0000-0001-9233-972X>

Daan Moechars

VIB KU Leuven

Sarah Borrie

VIB/KU Leuven Center for Brain & Disease Research <https://orcid.org/0000-0002-6880-772X>

An Snellinx

KU Leuven

Pierre Bourdely

VIB/KU Leuven Center for Cancer Biology

Nicola Thrupp

VIB/KU Leuven Center for Brain & Disease Research

Sebastiaan Moonen

VIB/KU Leuven Center for Brain & Disease Research

Emanuela Pasciuto

VIB/UAntwerpen Center for Molecular Neurology

Stephanie Humblet-Baron

KU Leuven

Joshua Stevenson-Hoare

Cardiff University

Emily Simmonds

UK DRI - Cardiff University

Dietmar Thal

University Hospitals Leuven <https://orcid.org/0000-0002-1036-1075>

Massimiliano Mazzone

VIB-KU Leuven <https://orcid.org/0000-0001-8824-4015>

Tracy Young-Pearse

Harvard Medical School <https://orcid.org/0000-0001-7846-1542>

Rik Vandenberghe

<https://orcid.org/0000-0001-6237-2502>

Valentina Escott-Price

Cardiff University <https://orcid.org/0000-0003-1784-5483>

Mark Fiers

VIB/KU Leuven Center for Brain & Disease Research

Biological Sciences - Article

Keywords:

Posted Date: November 10th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7794221/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: **Yes** there is potential Competing Interest. B.D.S. has been a consultant for Eli Lilly, Biogen, Janssen Pharmaceutica, Eisai, AbbVie and other companies and is now consultant to Muna Therapeutics. B.D.S is a scientific founder of Augustine Therapeutics and a scientific founder and stockholder of Muna Therapeutics. M.F. is a consultant to Muna Therapeutics. R.V.'s institution has clinical trial agreements (R.V. as PI) with Alector, AviadoBio, BMS, Eli Lilly, Johnson & Johnson, and UCB. R.V.'s institution has consultancy agreements (R.V. as DSMB/DMC member) with AC Immune and Novartis. D.R.T. received consultant fees from Muna Therapeutics and collaborated with Novartis Pharma AG and GE-Healthcare. All other authors have no competing interests.

Title: Polygenic risk for Alzheimer's disease shapes microglial inflammatory and antigen-presentation programs *in vivo*

Annerieke Sierksma^{1,2*†}, Daan Moechars^{1,2†}, Sarah C. Borrie^{1,2†}, An Snellinx^{1,2}, Pierre Bourdely^{3,4}, Nicola Thrupp^{1,2}, Sebastiaan Moonen^{1,2}, Emanuela Pasciuto⁵, Stephanie Humblet-Baron⁶, Joshua Stevenson-Hoare^{7,8}, Emily Simmonds⁷, Dietmar Rudolf Thal^{9,10}, Massimiliano Mazzone^{3,4}, Tracy L. Young-Pearse^{11,12}, Rik Vandenberghe^{13,14}, Valentina Escott-Price^{7,15}, Mark Fiers^{1,2,16,17*}, Bart De Strooper^{1,2,17,18*}

Affiliations

† These authors contributed equally to this work

¹ Department of Neurosciences and Leuven Brain Institute, KU Leuven; Leuven, Belgium

² VIB Center for Brain & Disease Research, VIB; Leuven, Belgium

³ Center for Cancer Biology, VIB; Leuven, Belgium

⁴ Department of Oncology, KU Leuven; Leuven, Belgium

⁵ VIB Center for Molecular Neurology, University of Antwerp; Antwerp, Belgium

⁶ Laboratory of Adaptive Immunology, Department of Microbiology, Immunology and Transplantation, KU Leuven; Leuven, Belgium

⁷ Dementia Research Institute, Cardiff University; Cardiff, UK

⁸ Current affiliations: Division of Clinical Epidemiology and Aging Research, German Cancer Research Center; Im Neuenheimer Feld 280, 69120 Heidelberg, Germany & Health and Life Science Alliance Heidelberg Mannheim; Berlinerstrasse 47, 69120 Heidelberg, Germany

⁹ Laboratory for Neuropathology, Department of Imaging and Pathology, Leuven Brain Institute, KU Leuven; Leuven, Belgium

¹⁰ Department of Pathology, University Hospitals Leuven; Leuven, Belgium

¹¹ Ann Romney Center for Neurologic Diseases, Department of Neurology, Brigham and Women's Hospital and Harvard Medical School; Boston, MA, USA

¹² Harvard Stem Cell Institute, Harvard University; Cambridge, MA, USA

¹³ Laboratory for Cognitive Neurology, Alzheimer Research Centre KU Leuven, Leuven Brain Institute; Leuven, Belgium

¹⁴ Neurology Department, University Hospitals Leuven; Leuven, Belgium

¹⁵ Division of Psychological Medicine and Clinical Neurosciences, Cardiff University; Cardiff, UK

¹⁶ Department of Human Genetics, KU Leuven; Leuven, Belgium

¹⁷ UK Dementia Research Institute at UCL, University College London; London, UK

¹⁸ Francis Crick Institute, London NW1 1AT, UK

*Corresponding authors: bart.destrooper@kuleuven.be, annerieke.sierksma@kuleuven.be, mark.fiers@kuleuven.be

Abstract

How polygenic risk translates into cellular dysfunction remains largely unknown in Alzheimer's disease and related disorders. Here, we selected 31 donors broadly spanning the Caucasian-based polygenic risk distribution for AD and show that human polygenic architecture modulates microglial immune responses to amyloid pathology *in vivo*. We developed a "microglia village" model by xenotransplanting pooled iPSC-derived microglia from these genetically diverse donors into amyloid-bearing (*App^{NLGF}*) and control (*App^{Hu}*) mouse brains, allowing the effect of genetic background to be separated from shared environmental influences. Marked inter-donor transcriptomic differences were observed in hMG derived from homeostatic, non-amyloid brain environments, demonstrating divergent baseline states across individuals. Amyloid exposure induced highly varied expression of MHC class II genes across donor, which correlated with individual's AD polygenic risk scores. These findings demonstrate that polygenic risk can be decoded into functional immune phenotypes in human microglia and establish a scalable *in vivo* platform to dissect the genetic regulation of cellular responses in complex brain disorders.

Introduction

Alzheimer's disease unfolds through a series of cellular tipping points, and genetic factors strongly influence how the brain navigates them. These include the build-up of amyloid pathology, microglial responses that seed or compact amyloid, the induction of pTAU and granulovacuolar pathology in neurons, vascular pathology, and the onset of neuronal loss¹⁻³. With high heritability estimates based on common SNP variants ranging from 0.38 to 0.66⁴, genetic risk is likely to determine an individual's relative vulnerability or resilience at each of these stages.

Genome-wide association studies (GWAS) have identified hundreds of common variants associated with AD^{5,6}, each conferring only modest individual risk (typically 1-3%, excluding APOE⁵). As understanding has shifted from an oligogenic⁷ to a polygenic view of AD^{4,8,9}, it is now evident that thousands of sub-threshold variants, together with rare coding and structural variants^{10,11}, collectively shape susceptibility. Protective genotypes, such as the APOE Christchurch mutation¹² and resilience-associated variants in centenarians¹³, further underscore the importance of studying the genome as an integrated whole rather than focusing on isolated loci. Yet, translating this complex genetic architecture into mechanistic insight remains a central challenge, due largely to the lack of models that can resolve genome-wide variation at cellular resolution¹⁴.

In addition to rare mutations causing familial AD and associated with amyloid accumulation¹⁵, a large fraction of common AD-associated variants (accounting for 69-84% of SNP-based heritability⁴), are enriched in genes expressed by microglia or located in microglia-specific enhancer regions^{5,15,16}. Microglia sense amyloid- β (A β) pathology, a key trigger of AD pathogenesis³, and initiate inflammatory cascades that influence neuronal health and function^{15,17-19}. In response to A β , human microglia (hMG) adopt transcriptional states characterized by cytokine signaling, antigen presentation (human leukocyte antigen (HLA)/major histocompatibility complex (MHC)) and disease-associated microglia (DAM)-like phenotypes^{20,21}. Many AD risk genes are differentially expressed across these states, suggesting that genetic variation tunes how microglia respond to pathological cues^{17,20,22}. This positions microglia as both mediators and interpreters of polygenic risk. Yet, direct evidence that natural human genetic diversity shapes microglial function *in vivo* remains lacking²³⁻²⁵.

To address this gap, we developed a pooled xenotransplantation model (the "microglia village") that enables high-resolution mapping of genotype-to-phenotype relationships *in vivo*²⁶. By engrafting iPSC-derived microglia from genetically diverse donors into the brains of amyloid-plaque-bearing²⁷ and control mice²⁸, we created a chimeric environment in which human microglia with different genotypes coexist and respond to identical cues within the same brain^{20,29-32}. This design isolates donor-intrinsic effects from the shared host environment, allowing systematic assessment of how polygenic risk shapes microglial responses to A β pathology. Using a relatively small but well-characterized iPSC cohort, we provide a first mechanistic demonstration of linking polygenic risk to functional cellular outcomes. More broadly, the microglia village provides a scalable, generalizable platform to dissect the functional impact of complex human genetics across neurological and psychiatric disorders.

Results

1. *In vivo* microglia villages disentangle genomic and environmental contributions to the microglial transcriptome.

To investigate how genetic variation in AD risk shapes hMG phenotypes *in vivo*, we computed polygenic risk scores (PRS) for AD from genome-wide SNP data (excluding the APOE locus, GWAS p-value threshold <0.05) across 3 cohorts⁸. Two AD case/control datasets (UZ Leuven, ROS/MAP³³) and the HipSci collection³⁴ were standardized to the Caucasian subset of the 1000 Genomes Project³⁵. Thirty-one individuals (Table 1) spanning -3 to +4 standard deviations of the Caucasian-based polygenic risk distribution for AD were selected and iPSCs lines that passed quality control for pluripotency, stemness and karyotype were obtained (see Methods). For each donor, iPSC-derived microglial precursors were generated and grouped into three pools of 12-13 donors (the ‘microglia villages’), with three donor lines shared across all pools as internal standards. Each village was independently differentiated and transplanted into neonatal amyloid-producing *App*^{NLGF} mice or *App*^{Hu}-control mice (both on *Rag2*^{-/-}, *Il2rg*^{-/-}, *hCSF1*^{KI} background, see Methods), thereby creating mixed-donor “microglia villages” (Fig.1A)²⁶⁻²⁸. This pooled design enables high-throughput, controlled comparisons across genotypes while minimizing the number of recipient animals. By exposing all donor-derived hMG to identical Aβ and brain microenvironments, the model isolates donor-intrinsic genetic effects from shared environmental influences.

Six months post-engraftment, hMG were isolated from the right brain hemisphere (excluding cerebellum) of recipient mice (n=4-6 per strain per village) and analyzed by single-cell RNA sequencing. Cells were demultiplexed by host mouse using CiteSeq antibody tagging³⁶, and by donor genotype, using Souporecell³⁷, based on donor-specific SNP profiles (see Methods). Two donors were excluded for low recovery, leaving 29 unique donor lines for downstream analysis.

Within each village, donor representation ranged from 27-35% for the most abundant line to 0.4-1.2% for the least, with reproducible engraftment across mice for each village and across both recipient strains (Fig.1B, Fig.S1A). Skews in donor contribution primarily impacted statistical power for detecting differentially expressed genes (DEGs), but even low-abundance lines yielded hundreds of DEGs (Fig.S1C). Strong correlations between shared anchor donor lines across villages (median 70%), demonstrated reproducibility across independent differentiations (Fig.S1D). Subsampling analyses confirmed that lower correlations ((e.g. *Kol2* in village 1) were attributable to low cell numbers (<500; Fig.S1D). Thus, although individual donor representation varied, reproducibility across anchor lines and subsampling analyses indicates that donor-specific signals are consistent and outweigh technical variation. We obtained in total an average of 8,107 high quality hMG transcriptomes per donor (range: 537-29,964) amounting to 235,106 cells after integration with Seurat and Harmony. A unified UMAP embedding of all three villages across all experiments is shown in Fig.1A and Fig.S1B.

To assess donor-genotype effects while avoiding pseudo-replication (i.e. treating each cell as an independent replicate³⁸), we applied a generalized linear mixed-effects model with host mouse as a random effect, correcting for sequencing batch, read depth, and mouse host sex. An in-

silico permutation test randomizing cell-to-donor assignments, but keeping all other proportions stable, confirmed type I error control, with at most one false-positive gene per comparison (Fig.S1C)³⁸. This statistical model thus provides a robust instrument to disentangle genetic contributions to human microglial function *in vivo*, both under homeostatic conditions and in the presence of amyloid- β pathology (see also Fig.S1E and Methods).

2. Female microglia exhibit a stronger inflammatory response than male microglia.

Post-hoc regression analysis revealed substantial donor sex-effects: female-derived hMG expressed higher levels of MHC class I & II, interferon, and cytokine genes (Fig.1C, Supplementary Text ST1, Fig.S2A, Tables S1-2), consistent with an intrinsically heightened immune activation potential. All subsequent analyses were therefore performed on sex-corrected donor values.

3. Genetic heterogeneity shapes the microglial transcriptional landscape under homeostatic conditions.

To assess donor effects in the absence of pathology, we analyzed hMG from *App^{Hu}* mice, comparing each donor's gene expression profile to the grand mean of all donors (i.e. the average of donor means, which mitigates differences in total cell number). Donor-specific UMAP projections and DEG analysis revealed substantial transcriptional diversity (mean: 1,973 DEGs per donor, range 426-3,841; Fig.2A-B; Fig.S1E). Gene co-expression analysis, using WGCNA³⁹, identified 13 modules with distinct biological functions (Fig.2C, Tables S3-4), ranging from discrete to very broad-range distributions across the UMAP (Fig.2D). Because modules can be expressed across multiple cell states, this approach better captures the multifunctional nature of microglia than cluster-based methods⁴⁰ (Fig.2C-D and Fig.S3), yet still encompasses previously identified transcriptional profiles such as interferon response microglia (Blue (Innate Immunity)) and specific cytokine responses (e.g. Yellow (Cytokine, chemokine and toll-like receptor signaling))²⁰. WGCNA confirmed that most donors diverged from the cohort average in homeostatic conditions in at least one gene module with over half of the modules showing more than 20 donors deviating significantly from the grand mean (Fig.2E-F, Fig.S7).

These findings confirm that genetic background alone primes microglia into distinct immune-competent states (Fig.2F). For example, most donors show Magenta (MHCII) module (Fig.2F) expression, containing MHCII antigen presentation program genes (e.g. *CD74*, *HLA-DRA*, *HLA-DRB* and the MHCII transactivator *CIITA*), below the amyloid-induced cohort response (red diamond, indicated as reference and based on the calculations detailed below), yet donors L233, L25 and L23 exhibit comparable or higher scores even under control conditions. The data challenge the notion of a uniform 'homeostatic' microglial state as inferred from xenograft models^{20,31} or post-mortem tissue^{21,41,42} and underscores the influence of genotype on baseline immune tone. These baseline differences provide a potential substrate through which polygenic

burden exerts effects, but larger donor cohorts will be needed to resolve the full spectrum of genetically encoded immune tones.

4. A consensus microglial response to amyloid includes a robust MHCII antigen presentation signature across donors.

An analysis of the previously characterized amyloid plaque response of microglia (ARM)^{20,43} across all donors provides a broader generalizability than the prior studies relying on one or two donor lines^{20,31}. To define the average human microglial response to A β , we compared transcriptomes from *App*^{NLGF} mice (with A β plaques) to those from *App*^{Hu} controls (lacking AD-related pathology) across all 29 genetically distinct donors ('strain effect' in Fig.S1E). We identified 2,262 differentially expressed genes (DEGs: 619 upregulated; 1,643 downregulated; Fig.3A, Table S5). GO and KEGG analysis (Table S6) revealed downregulation of RNA splicing, mRNA metabolism and histone modification and upregulation of MHCII related genes (*HLA-DRA*, *HLA-DPA1*, *HLA-DQA1*, *CD74*, *CIITA*) and huDAM markers (*CD9*, *APOC1*, *APOE*, *TREM2*, *MITF*, *MYO1E*, *LIPA*) consistent with prior studies^{20,21,31,41,42,44}.

In concordance, the Magenta (MHCII) and Brown (huDAM) WGCNA modules were upregulated, the latter module overlapping significantly with the huDAM state (e.g. *SPPI1*, *CD9*, and *GPNMB*; Fisher's exact test, odds ratio = 8.4, p=5.15e-15)²⁰. The Brown module is enriched for lysosomal, endocytosis, and lipid transport pathways (Fig.2B, Table S6). Additional upregulated modules are Turquoise (Ribosome) and Pink (Morphology & Immune) (Fig.2D). Together, these findings indicate that hMG mount a robust MHCII and huDAM-like response to A β -plaques confirming previous findings with one genotype²⁰, while downregulating transcriptional regulatory programs.

5. The MHCII response in microglia is functionally relevant and present in AD microglia in human brain.

We next confirmed at the protein level that plaque-exposed hMG upregulate antigen-presentation machinery. FACS analysis of hMG derived from H9 embryonic stem cells in *App*^{NLGF-FIRE} brains (a line known to induce MHCII expression upon amyloid exposure²⁰) showed increased surface expression of HLA-DR, HLA-ABC, and key co-stimulatory and -inhibitory molecules (CD80, CD86, PD-L1), compared to controls (Fig.2C-H). In both xenografted *App*^{NLGF} brains (Fig.S5A) and post-mortem AD tissue (Fig.S5B), MHCII expression was highest in microglia adjacent to plaques, consistent with literature^{20,45,46}. Thus, hMG exposed to amyloid plaques activate the full antigen-presenting machinery necessary for T cell engagement. Because xenotransplants were performed in *Rag2*-deficient mice lacking adaptive immunity, we next assessed the functional relevance of this phenotype ex vivo. hMG isolated from plaque-bearing mice stimulated naïve CD4⁺ and CD8⁺ T cells more effectively than controls, inducing cytokine production (TNF α , IFN γ and IL2) and T-cell proliferation (Fig.S6A-F). This confirms that amyloid exposure induces a bona fide antigen-presenting microglia state. Despite this capability,

T cell infiltration into the brain parenchyma appears limited in postmortem AD tissue (see Fig. S5C and⁴⁷) compared to other diseases like MS, suggesting that the role of MHCII⁺ microglia in AD and their interaction with T-cells requires further investigation.

6. Donor-specific microglial responses to A β pathology reveal striking variability in MHCII activation.

Across donors, hMG showed robust induction of MHCII and huDAM genes in response to amyloid plaques (Fig.2). In contrast, donor-level projections onto the UMAP revealed pronounced heterogeneity (Fig.4A), indicating substantial deviations from the consensus “average amyloid response”. DE analysis of the strain*donor interaction (Fig.S1E) confirmed marked inter-donor divergence in A β responses which was further supported by WGCNA showing distinct module activity profiles for each donor (Fig.4B-C, Fig.S7).

The Magenta (MHCII) module showed the most pronounced variability: 21 of 29 donors deviated significantly from the average across donors (Fig.4B-C, Fig.S7). For example, donor BR99 showed a two-fold increase in MHCII induction, whereas donor Oupf6 showed no detectable induction (Fig.4C). Donor variation also influenced the Brown (huDAM) and Pink (morphology) modules (9 and 11 donors respectively) (Fig.4C, Fig.S7), as well as modules not strongly activated at the cohort-level average such as Blue (Innate immunity) and Green (Endoplasmic reticulum). Together these findings illustrate that amyloid plaques elicit a conserved ARM response in human microglia, even across genetically diverse donors, while also revealing extensive donor-to-donor variability in the magnitude and composition of this response. The Magenta (MHCII) module emerges as the most genetically sensitive component of the A β -induced microglial program. Importantly, these results do not imply that MHCII is the sole relevant pathway, but rather that it represents the most statistically tractable and reproducible signal given the current sample size. To ensure that these donor-specific effects were not artifacts of the pooled design, we next validated transcriptional concordance in independently differentiated, single-donor xenografts.

7. Single-donor grafts confirm village-based transcriptional profiles

To validate that village-derived hMG preserve donor-specific traits, we transplanted 16 iPSC lines individually into *App*^{NLGF} mice and compared their transcriptomes to those of the same donors in microglia villages. Gene-level concordance was high (average Pearson's $r = 0.81$; Fig.S8A). This concordance suggests that the pooled design does not artificially generate phenotypic diversity but rather captures reproducible donor-intrinsic traits. The results demonstrate that village-based xenotransplantation faithfully preserves inherent, donor-specific transcriptional responses of hMG across differentiation batches, supporting its use as a scalable and robust platform to capture cell-autonomous effects of genetic variation *in vivo*.

8. Polygenic risk predicts MHCII responses to amyloid pathology

To retain statistical power given our modest sample size ($n=29$ donors), we limited the number of hypotheses tested. Based on the marked donor-to-donor variability in MHCII responses both to plaques and at baseline, we focused on whether clinical diagnosis or polygenic risk (PRS) for AD⁸, explained variation in Magenta (MHCII) module expression. Among donors with known diagnosis (AD cases, $n=13$; controls, $n=11$), we observed no baseline differences ($t(23)=0.05$, $p=0.96$) but plaque-induced MHCII responses were significantly stronger in AD cases ($t(23)=3.53$, $p=0.0019$; Fig.5A). Gene-level analysis confirmed that this effect was driven by the eleven MHCII-related genes in the module, rather than by the two MS4A-locus genes (Fig.5B).

Whole-genome PRS (excluding the APOE locus and $pT<0.05$ based on⁶) correlated strongly with donor-specific Magenta (MHCII) induction in response to A β ($R=0.71$, $p=9.6E-5$, Fig.5C). A similar, but non-significant, correlation was observed in the hMG derived from the Hipsci subset for which we have no diagnostic data ($R=0.71$, $p=0.17$, Fig.S9A). Using partial correlation analysis we confirmed that the association between the donor-specific Magenta response to A β in AD cases and controls and PRS was not driven by confounders such as the APOE genotypes in our cohort or donor proportions in our villages and retained its significance (after controlling for APOE: $R=0.65$, $p=0.0005$; after controlling for donor proportion: $p=0.70$, $p=0.0001$). Moreover, we could not identify enrichment of AD-associated MHCII haplotypes among MHCII-high donors (Table S7). No correlation was detected between PRS and baseline MHCII expression (Fig.S9B). Thus, the correlation between PRS and MHCII responses to A β appears polygenic in nature and is not explained by APOE, donor distributions within the village or single HLA haplotypes, although future studies with larger sample sizes will be required to resolve contributions from specific loci. In conclusion, these findings demonstrate that polygenic AD risk predicts the strength of microglial MHCII activation in response to amyloid pathology, supporting a model in which polygenic risk burden shapes stimulus-specific immune activation in human microglia.

Discussion

Genes implicated by AD-associated GWAS loci likely act across multiple cell types and disease mechanisms, including plaque pathology formation and neuronal vulnerability to protein aggregation. The enrichment of AD risk loci in microglial genes and enhancers^{16,48,49}, suggests that part of the heritable predisposition to AD operates through altered microglial responses to pathology. Yet, in sporadic AD, where risk is largely polygenic and distributed genome-wide, it has remained challenging to translate PRS into concrete biological mechanisms¹⁵. Here, we provide *in vivo* functional evidence that elements of the polygenic architecture can be resolved into stimulus-specific cellular read-outs where polygenic burden correlates with the strength of microglial immune responses to amyloid pathology. Remarkably this was detectable in a relatively small but carefully characterized iPSC cohort, suggesting that functional decoding of PRS may be feasible even at moderate cohort sizes. This does not preclude that other elements of polygenic risk act on neurons, vasculature, or other glia populations, but it establishes for the first time that PRS can be linked directly to measurable immune phenotypes *in vivo*.

By leveraging genetic diversity in a pool of xenografted human iPSC-derived microglia, we isolate donor-intrinsic transcriptional variation from environmental influences. Our findings show that human genomic variation shapes both baseline and amyloid-induced microglial states and that polygenic risk for AD primarily tunes the magnitude of microglial reactivity rather than the baseline state. Notably, microglia from some donors exhibited elevated expression of MHC class II and human disease-associated microglia (huDAM) genes even in the absence of pathology, challenging the notion of a uniform “homeostatic” state^{21,41,42,44}.

Across donors, microglia responses to A β pathology converged on a shared transcriptional program marked by inflammatory and antigen presentation genes, consistent with previous observations in AD tissue and models^{20,21,31,41,42,44}. However, donor-specific analysis revealed striking variability which may contribute to explaining inconsistencies between different postmortem AD datasets^{21,41,42,44}, and highlights why future studies will require even larger donor cohorts to map the full landscape of genetically encoded microglial states. Importantly, sex contributed significantly to this diversity: female-derived microglia exhibited stronger induction of MHCII and cytokine pathways, suggesting that genetic and sex-linked factors jointly modulate immune activation thresholds. This observation aligns with epidemiological and molecular evidence that sex modifies AD risk and progression^{50,51} and it underscores the need to consider sex as an integral dimension of microglial heterogeneity.

Our analysis suggests that module-based frameworks such as WGCNA which capture overlapping functional programs, may provide a more robust and transferable approach for classifying microglia states across disease contexts. Among these programs, antigen presentation via MHCII emerged as the strongest correlate of polygenic risk in this current study. Nevertheless, other pathways, including interferon signaling, lipid metabolism and cytokine networks, also varied across donors, suggesting that multiple immune axes may be genetically tuned and should be explored in expanded cohorts.

The data confirm that the antigen-presentation program is both reproducible and functionally relevant. Although the immunocompromised host strain may influence the cytokine milieu, the strong concordance with human post-mortem data showing HLA-DR positive microglia around amyloid plaques (Fig.S5B) supports the physiological relevance of these findings. Amyloid exposure induced MHCII and key co-stimulatory and inhibitory molecules (CD80, PD-L1), shifting hMG toward an antigen-presenting state capable of enhanced T cell activation^{52–54}. Crucially, the magnitude of this response varied across donors and correlated with their AD polygenic scores, highlighting its potential role in disease susceptibility^{20,31,44–46}. While specific HLA haplotypes have been linked to AD^{5,6,55}, we found that risk-associated and protective HLA variants occurred across the full spectrum of MHCII module expression, rather than single-allele effects. Similarly, hMGs carrying rare variants such as TREM2-R47H, PLCG2-P522R or SORL1 knockout, also showed altered MHCII expression^{20,56,57}. Together these data raise the possibility that individuals with high genetic burden may exhibit exaggerated immune responses to pathology, a concept that could guide stratified immunomodulatory therapies.

Our data do not yet establish whether MHCII-high microglia are protective, maladaptive, or compensatory in AD; future studies incorporating neuronal readouts and adaptive immunity

will be essential to define these functional consequences. In our xenograft model, MHCII upregulation occurs in the absence of peripheral immune cells, owing to the immunocompromised host background, demonstrating that this response can arise cell-autonomously in microglia⁵⁸. In AD, T-cells are relatively scarce in the parenchyma (Fig.S5C)⁴⁷, in contrast to multiple sclerosis, where immune infiltration is pronounced⁵⁹. This raises the question of whether MHCII+ microglia represent a maladaptive state in AD. It remains plausible that MHCII-high or -low microglia modulate the CNS environment or contribute to neuronal dysfunction in addition to direct T cell interactions.

The hMG village model provides a powerful framework for in vivo functional polygenomics, enabling dissection of how complex genetic backgrounds shape microglial immune phenotypes. The strong transcriptional concordance between pooled and single-donor grafts validates its scalability. Accounting for sex as a biological variable proved essential, both to isolate donor-intrinsic effects and to reveal sex-linked amplification of immune pathways⁶⁰. This further supports the view that microglial activation thresholds and transcriptional programs are set by an interplay between genetic background, sex, and pathological cues^{50,51}.

Conceptually, our findings begin to recast polygenic risk from a statistical abstraction into a cell-intrinsic modifier of immune function. By explicitly demonstrating this in vivo, we provide a framework that can be generalized to larger-scale cohorts, different microglial stimuli, and other cell types implicated by AD genetics. Microglial activation in the current work emerges not as a uniform response to pathology, but as a genetically encoded trait that varies across individuals even under identical conditions. This framework enables genotype- and sex-stratified prediction of microglial responses to pathology and establishes a tractable platform for personalized pharmacogenomic screening. Beyond AD, similar complex genetic architectures may shape stimulus-specific microglial behaviors in psychiatric, neurodevelopmental, and other neurodegenerative disorders where microglia play increasingly recognized roles⁶¹.

References

1. Baligács, N. *et al.* Homeostatic microglia initially seed and activated microglia later reshape amyloid plaques in Alzheimer's Disease. *Nat Commun* **15**, 10634 (2024).
2. Balusu, S. & De Strooper, B. The necroptosis cell death pathway drives neurodegeneration in Alzheimer's disease. *Acta Neuropathol* **147**, 96 (2024).
3. Karran, E. & De Strooper, B. The amyloid hypothesis in Alzheimer disease: new insights from new therapeutics. *Nat Rev Drug Discov* **21**, 306–318 (2022).
4. Baker, E. *et al.* What does heritability of Alzheimer's disease represent? *PLOS ONE* **18**, e0281440 (2023).
5. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet* **54**, 412–436 (2022).
6. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat Genet* **51**, 414–430 (2019).
7. Zhang, Q. *et al.* Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. *Nat Commun* **11**, 4799 (2020).
8. Leonenko, G. *et al.* Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores. *Nat Commun* **12**, 4506 (2021).
9. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177–1186 (2017).
10. Wang, H., Wang, L.-S., Schellenberg, G. D. & Lee, W.-P. The role of structural variations in Alzheimer's disease and other neurodegenerative diseases. *Frontiers in Aging Neuroscience* **14**, (2023).
11. De Deyn, L. & Sleegers, K. The impact of rare genetic variants on Alzheimer disease. *Nat Rev Neurol* **21**, 127–139 (2025).

12. Quiroz, Y. T. *et al.* APOE3 Christchurch Heterozygosity and Autosomal Dominant Alzheimer's Disease. *New England Journal of Medicine* **390**, 2156–2164 (2024).
13. Tesi, N. *et al.* Cognitively healthy centenarians are genetically protected against Alzheimer's disease. *Alzheimer's & Dementia* **20**, 3864–3875 (2024).
14. Flint, J. & Ideker, T. The great hairball gambit. *PLOS Genetics* **15**, e1008519 (2019).
15. Sierksma, A., Escott-Price, V. & De Strooper, B. Translating genetic risk of Alzheimer's disease into mechanistic insight and drug targets. *Science* **370**, 61–66 (2020).
16. Nott, A. *et al.* Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science* **366**, 1134–1139 (2019).
17. Sierksma, A. *et al.* Novel Alzheimer risk genes determine the microglia response to amyloid- β but not to TAU pathology. *EMBO Mol Med* **12**, e10606 (2020).
18. Sims, R., Hill, M. & Williams, J. The multiplex model of the genetics of Alzheimer's disease. *Nat Neurosci* **23**, 311–322 (2020).
19. De Schepper, S., Crowley, G. & Hong, S. Understanding microglial diversity and implications for neuronal function in health and disease. *Developmental Neurobiology* **81**, 507–523 (2021).
20. Mancuso, R. *et al.* Xenografted human microglia display diverse transcriptomic states in response to Alzheimer's disease-related amyloid- β pathology. *Nat Neurosci* **27**, 886–900 (2024).
21. Sun, N. *et al.* Human microglial state dynamics in Alzheimer's disease progression. *Cell* **186**, 4386–4403.e29 (2023).
22. Sims, R. *et al.* Rare coding variants in PLCG2, ABI3 and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. *Nat Genet* **49**, 1373–1384 (2017).
23. Yang, H. S. *et al.* Natural genetic variation determines microglia heterogeneity in wild-derived mouse models of Alzheimer's disease. *Cell Reports* **34**, (2021).

24. Neuner, S. M., Heuer, S. E., Huentelman, M. J., O'Connell, K. M. S. & Kaczorowski, C. C. Harnessing Genetic Complexity to Enhance Translatability of Alzheimer's Disease Mouse Models: A Path toward Precision Medicine. *Neuron* **101**, 399-411.e5 (2019).
25. Antón-Bolaños, N. *et al.* Brain Chimeroids reveal individual susceptibility to neurotoxic triggers. *Nature* **631**, 142–149 (2024).
26. Wells, M. F. *et al.* Natural variation in gene expression and viral susceptibility revealed by neural progenitor cell villages. *Cell Stem Cell* **30**, 312-332.e13 (2023).
27. Saito, T. *et al.* Single App knock-in mouse models of Alzheimer's disease. *Nat Neurosci* **17**, 661–663 (2014).
28. Serneels, L. *et al.* Modeling the β -secretase cleavage site and humanizing amyloid-beta precursor protein in rat and mouse to study Alzheimer's disease. *Molecular Neurodegeneration* **15**, 60 (2020).
29. Fattorelli, N. *et al.* Stem-cell-derived human microglia transplanted into mouse brain to study human disease. *Nat Protoc* **16**, 1013–1033 (2021).
30. Xu, R. *et al.* Human iPSC-derived mature microglia retain their identity and functionally integrate in the chimeric mouse brain. *Nat Commun* **11**, 1577 (2020).
31. Hasselmann, J. *et al.* Development of a Chimeric Model to Study and Manipulate Human Microglia In Vivo. *Neuron* **103**, 1016-1033.e10 (2019).
32. Bohlen, C. J. *et al.* Diverse Requirements for Microglial Survival, Specification, and Function Revealed by Defined-Medium Cultures. *Neuron* **94**, 759-773.e8 (2017).
33. Lagomarsino, V. N. *et al.* Stem cell-derived neurons reflect features of protein networks, neuropathology, and cognitive outcome of their aged human donors. *Neuron* **109**, 3402-3420.e9 (2021).

34. Streeter, I. *et al.* The human-induced pluripotent stem cell initiative—data resources for cellular genetics. *Nucleic Acids Research* **45**, D691–D697 (2017).
35. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
36. Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biology* **19**, 224 (2018).
37. Heaton, H. *et al.* Souporecell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat Methods* **17**, 615–620 (2020).
38. Squair, J. W. *et al.* Confronting false discoveries in single-cell differential expression. *Nat Commun* **12**, 5692 (2021).
39. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Reports Methods* **3**, (2023).
40. Marshe, V. S. *et al.* A factor-based analysis of individual human microglia uncovers regulators of an Alzheimer-related transcriptional signature. 2025.03.27.641500 Preprint at <https://doi.org/10.1101/2025.03.27.641500> (2025).
41. Gabitto, M. I. *et al.* Integrated multimodal cell atlas of Alzheimer’s disease. *Nat Neurosci* **27**, 2366–2383 (2024).
42. Gazestani, V. *et al.* Early Alzheimer’s disease pathology in human cortex involves transient cell states. *Cell* **186**, 4438–4453.e23 (2023).
43. Sala Frigerio, C. *et al.* The Major Risk Factors for Alzheimer’s Disease: Age, Sex, and Genes Modulate the Microglia Response to A β Plaques. *Cell Reports* **27**, 1293–1306.e6 (2019).
44. Green, G. S. *et al.* Cellular communities reveal trajectories of brain ageing and Alzheimer’s disease. *Nature* 1–12 (2024) doi:10.1038/s41586-024-07871-6.

45. McGeer, P. L., Itagaki, S., Boyes, B. E. & McGeer, E. G. Reactive microglia are positive for HLA-DR in the substantia nigra of Parkinson's and Alzheimer's disease brains. *Neurology* **38**, 1285–1291 (1988).
46. Perlmuter, L. S., Scott, S. A., Barrón, E. & Chui, H. C. MHC class II-positive microglia in human brain: Association with alzheimer lesions. *Journal of Neuroscience Research* **33**, 549–558 (1992).
47. Togo, T. *et al.* Occurrence of T cells in the brain of Alzheimer's disease and other neurological diseases. *Journal of Neuroimmunology* **124**, 83–92 (2002).
48. Novikova, G. *et al.* Integration of Alzheimer's disease genetics and myeloid genomics identifies disease risk regulatory elements and genes. *Nat Commun* **12**, 1610 (2021).
49. Gjoneska, E. *et al.* Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* **518**, 365–369 (2015).
50. Beam, C. R. *et al.* Differences Between Women and Men in Incidence Rates of Dementia and Alzheimer's Disease. *J Alzheimers Dis* **64**, 1077–1083 (2018).
51. Coughlan, G. T. *et al.* Sex Differences in Longitudinal Tau-PET in Preclinical Alzheimer Disease: A Meta-Analysis. *JAMA Neurol* e250013 (2025) doi:10.1001/jamaneurol.2025.0013.
52. Lodygin, D. *et al.* A combination of fluorescent NFAT and H2B sensors uncovers dynamics of T cell activation in real time during CNS autoimmunity. *Nat Med* **19**, 784–790 (2013).
53. Chen, D. *et al.* CTLA-4 blockade induces a microglia-Th1 cell partnership that stimulates microglia phagocytosis and anti-tumor function in glioblastoma. *Immunity* **56**, 2086-2104.e8 (2023).
54. Chen, X. *et al.* Microglia-mediated T cell infiltration drives neurodegeneration in tauopathy. *Nature* **615**, 668–677 (2023).

55. Le Guen, Y. *et al.* Multiancestry analysis of the HLA locus in Alzheimer's and Parkinson's diseases uncovers a shared adaptive immune response mediated by HLA-DRB1*04 subtypes. *Proceedings of the National Academy of Sciences* **120**, e2302720120 (2023).
56. Claes, C. *et al.* The P522R protective variant of PLCG2 promotes the expression of antigen presentation genes by human microglia in an Alzheimer's disease mouse model. *Alzheimer's & Dementia* **18**, 1765–1778 (2022).
57. Ovesen, P. L. *et al.* Alzheimer's Disease Risk Gene SORL1 Promotes Receptiveness of Human Microglia to Pro-Inflammatory Stimuli. *Glia* **73**, 857–872 (2025).
58. Mittal, K. *et al.* CD4 T Cells Induce A Subset of MHCII-Expressing Microglia that Attenuates Alzheimer Pathology. *iScience* **16**, 298–311 (2019).
59. Neumann, H., Medana, I. M., Bauer, J. & Lassmann, H. Cytotoxic T lymphocytes in autoimmune and degenerative CNS diseases. *Trends in Neurosciences* **25**, 313–319 (2002).
60. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat Rev Immunol* **16**, 626–638 (2016).
61. Mondelli, V., Vernon, A. C., Turkheimer, F., Dazzan, P. & Pariante, C. M. Brain microglia in psychiatric disorders. *The Lancet Psychiatry* **4**, 563–572 (2017).
62. Humblet-Baron, S. *et al.* A systems immunology analysis of Alzheimer's disease reveals an age- and environmental exposure-independent disturbance in B cell maturation. 2025.06.16.25329692 Preprint at <https://doi.org/10.1101/2025.06.16.25329692> (2025).
63. Luckett, E. S. *et al.* Association of Alzheimer's disease polygenic risk scores with amyloid accumulation in cognitively intact older adults. *Alzheimer's Research & Therapy* **14**, 138 (2022).
64. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, s13742-015-0047–8 (2015).

65. Okita, K. *et al.* A more efficient method to generate integration-free human iPS cells. *Nat Methods* **8**, 409–412 (2011).
66. Marsh, S. E. *et al.* Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nat Neurosci* **25**, 306–316 (2022).
67. De Rop, F. V. *et al.* Hydrop enables droplet-based single-cell ATAC-seq and single-cell RNA-seq using dissolvable hydrogel beads. *eLife* **11**, e73971 (2022).
68. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* **42**, 293–304 (2024).
69. Hippen, A. A. *et al.* miQC: An adaptive probabilistic framework for quality control of single-cell RNA-sequencing data. *PLOS Computational Biology* **17**, e1009290 (2021).
70. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology* **20**, 296 (2019).
71. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296 (2019).
72. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
73. Brooks, M. E. *et al.* glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal* **9**, 378–400 (2017).
74. Lenth, Russell V. emmeans: Estimated Marginal Means, aka Least-Squares Means. (2025).
75. Korotkevich, G. *et al.* Fast gene set enrichment analysis. 060012 Preprint at <https://doi.org/10.1101/060012> (2021).
76. Moonen, S. *et al.* Pyroptosis in Alzheimer’s disease: cell type-specific activation in microglia, astrocytes and neurons. *Acta Neuropathol* **145**, 175–195 (2023).

- 521 77. García-Cabezas, M. Á., John, Y. J., Barbas, H. & Zikopoulos, B. Distinction of Neurons, Glia
522 and Endothelial Cells in the Cerebral Cortex: An Algorithm Based on Cytological Features.
523 *Front Neuroanat* **10**, 107 (2016).
- 524 78. Rojo, R. *et al.* Deletion of a Csf1r enhancer selectively impacts CSF1R expression and
525 development of tissue macrophage populations. *Nat Commun* **10**, 3215 (2019).
- 526
- 527

Acknowledgments

We warmly thank Veronique Hendrickx and Amber Claes for animal husbandry and genotyping; Leen Wolfs for single cell library preparations, Suresh Poovathingal, Marta Wojno and Niels Vandamme from the VIB Single Cell Core and the Single Cell Microfluidics & Analytics and Bioinformatics unit for help with library preparations and sequencing. Graphical schemes were created with BioRender.com.

Funding: This project received funding from:

European Research Council (ERC), ERC-834682 “CELLPHASE_AD” (BDS)
EU Joint Program – Neurodegenerative Disease Research, JPND2019-466-083, “TRIAGE” (BDS, VEP, RV)
VLAIO ICON grant for Personalized Medicine, HBC.2019.2523 “PRISMA” (BDS, VEP, RV)
Alzheimer’s Association USA, ADSF-23-1174466-C (BDS)
Alzheimer’s Association USA, 22-AAIA-963171 (DRT)
KU Leuven & Flemish Government, Methusalem METH/21/05 (BDS)
KU Leuven Opening the Future & Mission Lucidity (BDS)
Fonds Verhelst (BDS)
Flanders Institute for Biotechnology (VIB vzw) (BDS, MM, EP)
Fonds voor Wetenschappelijk Onderzoek (BDS)
Fonds voor Wetenschappelijk Onderzoek, G065721N, G024925N (DRT)
Queen Elisabeth Medical Foundation for Neurosciences (BDS)
UK Dementia Research Institute [award number UK DRI-1004] (BDS) which receives its funding from UK DRI Ltd, funded by the UK Medical Research Council
Alzheimer’s Society (BDS)
Alzheimer’s Research UK (BDS)
Medical Research Grant, MR/Y014847/1 (BDS)
Bax-Vanluffelen Chair for Alzheimer’s Disease (BDS)
Alzheimer Drug Discovery Foundation (BDS)

Author contributions:

Conceptualization: ASi, MF, BDS
Investigation & methodology: ASi, DM, SB, ASn, NT, PB, EP, MM, SM, DRT, JS-H, ES, RV, VEP, DRT

560 Software & visualization: ASi, DM, MF, ES, VEP
561 Writing – original draft: ASi, DM, BDS
562 Writing – review and editing: all authors
563 Resources: SH-B, RV, TLY-P, DRT
564 Supervision: ASi, MF, MM, DRT, VEP, RV, BDS
565 Funding acquisition: ASi, RV, VEP, BDS
566

567 **Competing interests:**

568 B.D.S. has been a consultant for Eli Lilly, Biogen, Janssen Pharmaceutica, Eisai, AbbVie and other
569 companies and is now consultant to Muna Therapeutics. B.D.S is a scientific founder of Augustine
570 Therapeutics and a scientific founder and stockholder of Muna Therapeutics. M.F. is a consultant
571 to Muna Therapeutics. R.V.’s institution has clinical trial agreements (R.V. as PI) with Alector,
572 AviadoBio, BMS, Eli Lilly, Johnson & Johnson, and UCB. R.V.’s institution has consultancy
573 agreements (R.V. as DSMB/DMC member) with AC Immune and Novartis. D.R.T. received
574 consultant fees from Muna Therapeutics and collaborated with Novartis Pharma AG and GE-
575 Healthcare. All other authors have no competing interests.

576 **Data and materials availability:**

577 Requests for further information and resources should be directed to and will be fulfilled by the
578 lead contact, Bart De Strooper (bart.destrooper@kuleuven.be). The raw read data will be deposited
579 in the EGA database under accession number [INSERT] and made accessible behind a data access
580 committee. The data can be viewed at (temporary website for review: <https://prs.nevel.eu>;
581 username: prs; password: XEK~ic%Q3hZh).
582

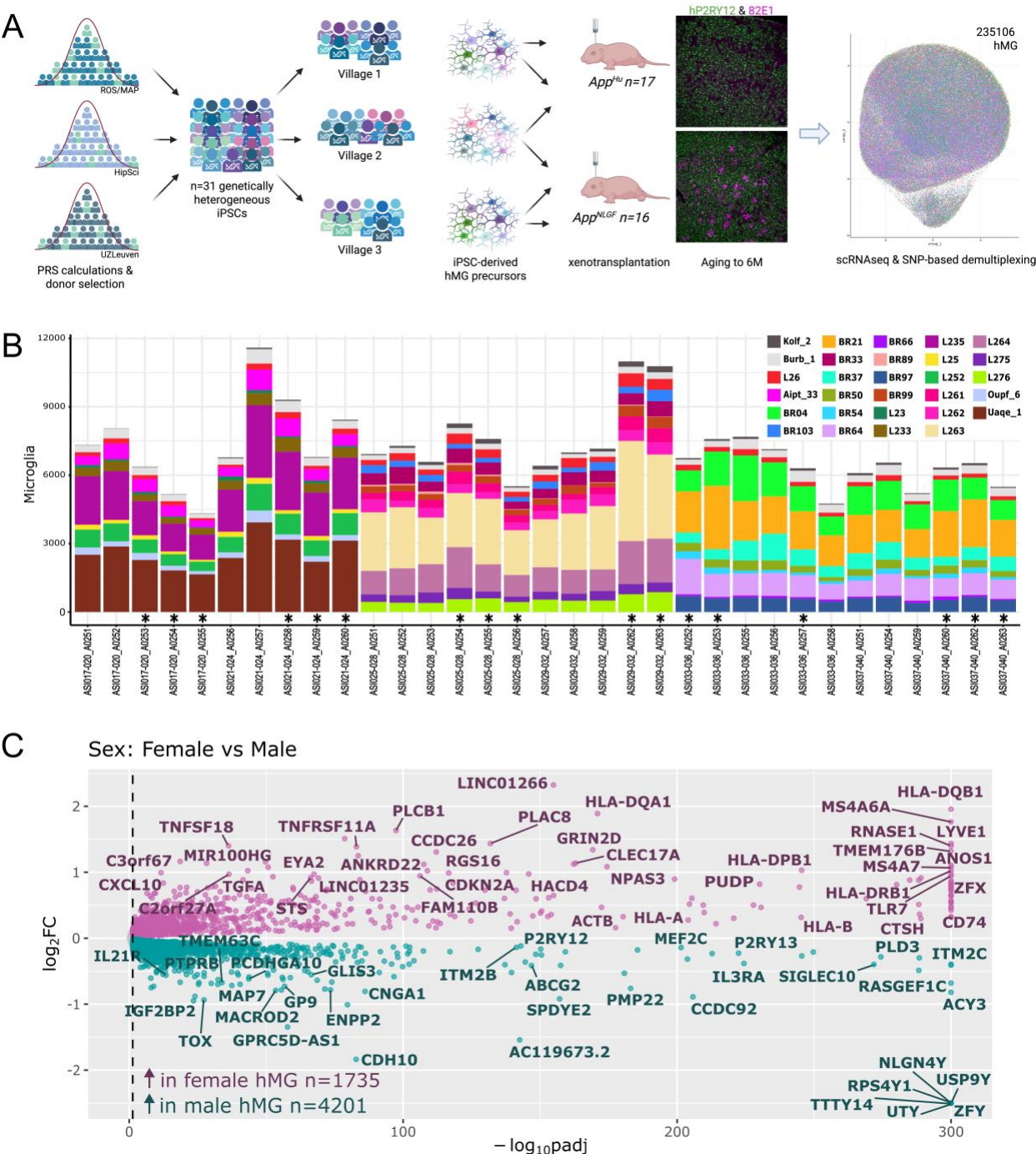


Fig. 1. The *in vivo* microglia village approach allows to robustly separate impact of genome and environment on the human microglial transcriptome

(A) Whole genome polygenic risk scores (PRS, excl the APOE locus, $pT<0.05$) were calculated for three cohorts of individuals (see Methods), and iPSCs from donors spanning a range of PRS

were obtained (represented as green individuals in each cohort). Human iPSC-derived microglial precursors from 31 genetically distinct donors were grouped into 3 pooled “villages” (n=12-13 lines per pool, with three lines shared across all villages). Each village was xenotransplanted at postnatal day 4 into 4-6 *App^{Hu}* (Control) and *App^{NLGF}* (amyloid producing) mice (n=17 and n=16 total, respectively, both on *Rag2^{-/-}*, *Il2rg^{-/-}*, *hCSF1^{KI}* background). At 6 months of age, human microglia (hMG) were isolated and profiled by single-cell RNA sequencing. A total of 235,106 high-quality hMG were assigned to individual host mice using Cite-seq hashing and to individual donors using SNP-based demultiplexing. Data from all donors and mice were integrated into a unified UMAP for downstream analysis.

(B) Stacked bar-graph of the number of hMG after demultiplexing per mouse (x-axis; *App^{NLGF}* mice are indicated with an asterisk) and per cell line (different colors). The three lines that were included in all 3 villages to serve as internal standards are at the top of each bar-graph (*Kol2* (dark grey), *Burbl* (light grey) and *L26* (red)), which show high reproducibility across different experiments (also see Fig.S1).

(C) A rotated volcano plot showing differentially expressed genes (DEGs) between hMG from female versus male donors. DEGs upregulated in female microglia (n=1735) are shown in purple; those upregulated in male microglia (n=4201) are in green (Benjamini–Hochberg–adjusted $p < 0.05$). Values are capped at $LFC < |2.5|$ and $-\log_{10}(padj) < 300$.

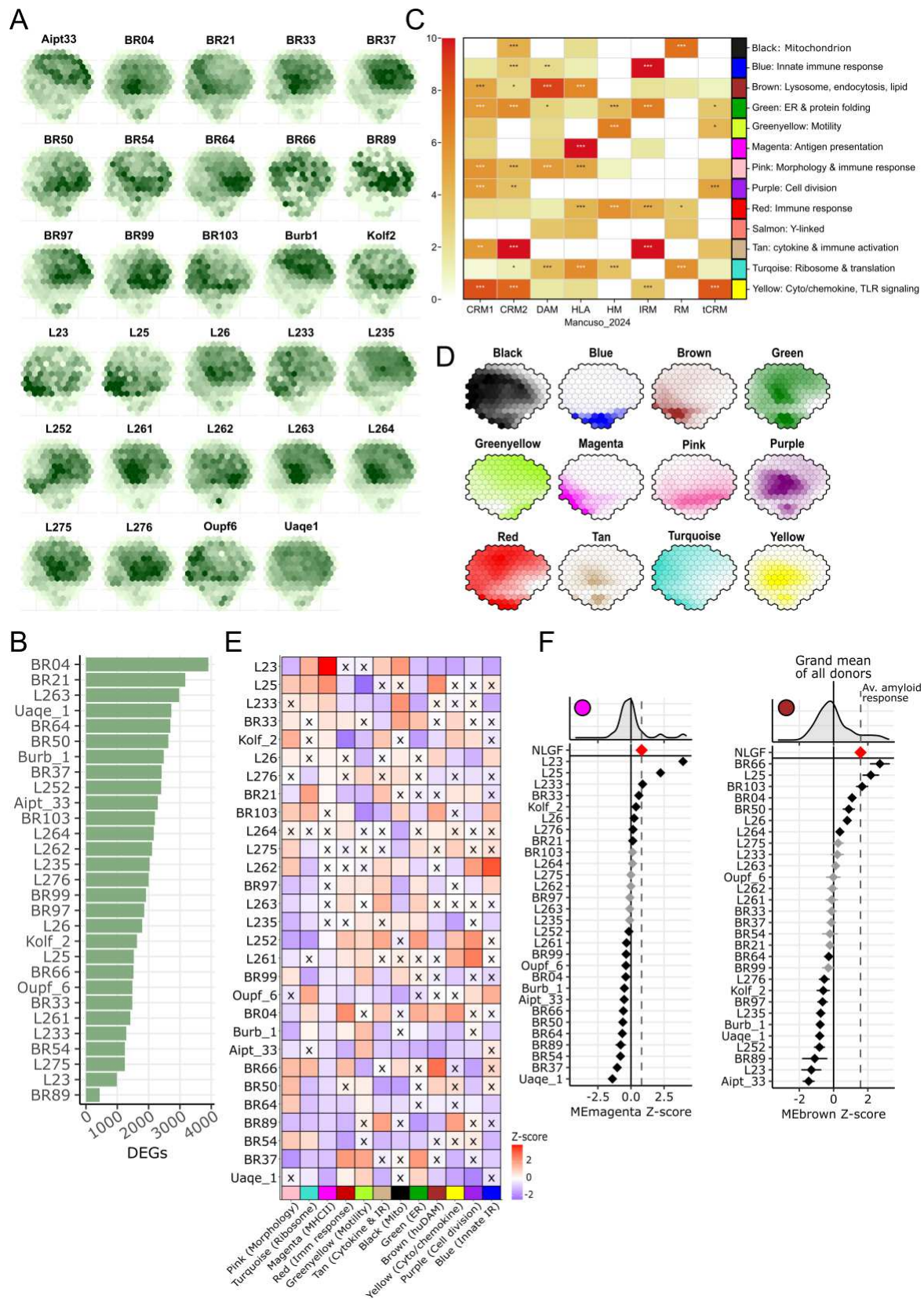


Fig. 2. Microglial transcriptomic diversity reflects genotype-specific homeostatic states in a physiological brain environment

(A) UMAP density plots showing the distribution of hMG from each donor in *App^{Hu}* mice, projected onto the integrated transcriptional space. Darker colors indicate where the cells of the indicated donor are distributed across transcriptional space.

(B) Bar-graph depicting the number of significant DEGs per donor in *App^{Hu}* mice compared to the average hMG profile across all donors.

(C) Overlap between the WGCNA co-expression modules and microglial transcriptional states (as defined in Mancuso et al.²⁰), with color scale displaying the log(odds ratio), capped at 10 and showing only enrichment. Modules (y-axis) were functionally annotated using GO & KEGG pathway enrichment. Cell states (x-axis) include: CRM1 (cytokine response microglia 1); CRM2 (cytokine response microglia 2); DAM (disease-associated microglia); HLA (antigen presenting microglia); HM (homeostatic microglia); IRM (interferon response microglia), RM (ribosomal microglia); tCRM (transitioning CRM). Other abbreviations: ER (endoplasmic reticulum); TLR (Toll-like receptor). *: Benjamini-Hochberg corrected p-value (padj)<0.05; **: padj<0.01; ***: padj<0.001.

(D) Spatial distribution of the WGCNA modules scores across the UMAP visualized using hexbin-based density plots. Darker UMAP color represents a higher average WGCNA enrichment score across the hMG within that hexbin. Certain modules occupy distinct transcriptional niches (e.g., Brown (huDAM); Magenta (MHCII)) while others show continuous gradient-like expression (e.g., Greenyellow (Motility); Green (ER & protein folding)).

(E) Module differential expression analysis for each donor. Deviations from the average transcriptional profile across all *App^{Hu}* microglia are indicated by colour. The x-axis lists WGCNA modules; the y-axis lists individual donors. Cells marked with “x” denote no significant difference from the average transcriptomic profile (BH-adjusted $p < 0.05$).

(F) Donor-specific z-score shifts in the Magenta (MHCII, left) and Brown (huDAM, right) modules in hMG from *App^{Hu}* mice. Each dot represents the mean module score across all cells from a given donor (error bars: 95% confidence interval), plotted relative to the average *App^{Hu}* expression across all donors (black vertical line at $x = 0$). The average *App^{NLGF}* response from Fig.3 is included for reference (top row, dashed line, listed as “NLGF”). Black diamonds indicate significant deviation from the *App^{Hu}* average (BH-adjusted $p < 0.05$); grey diamonds: not significant. Top density plots show the overall distribution of module z-scores across all *App^{Hu}* hMG.

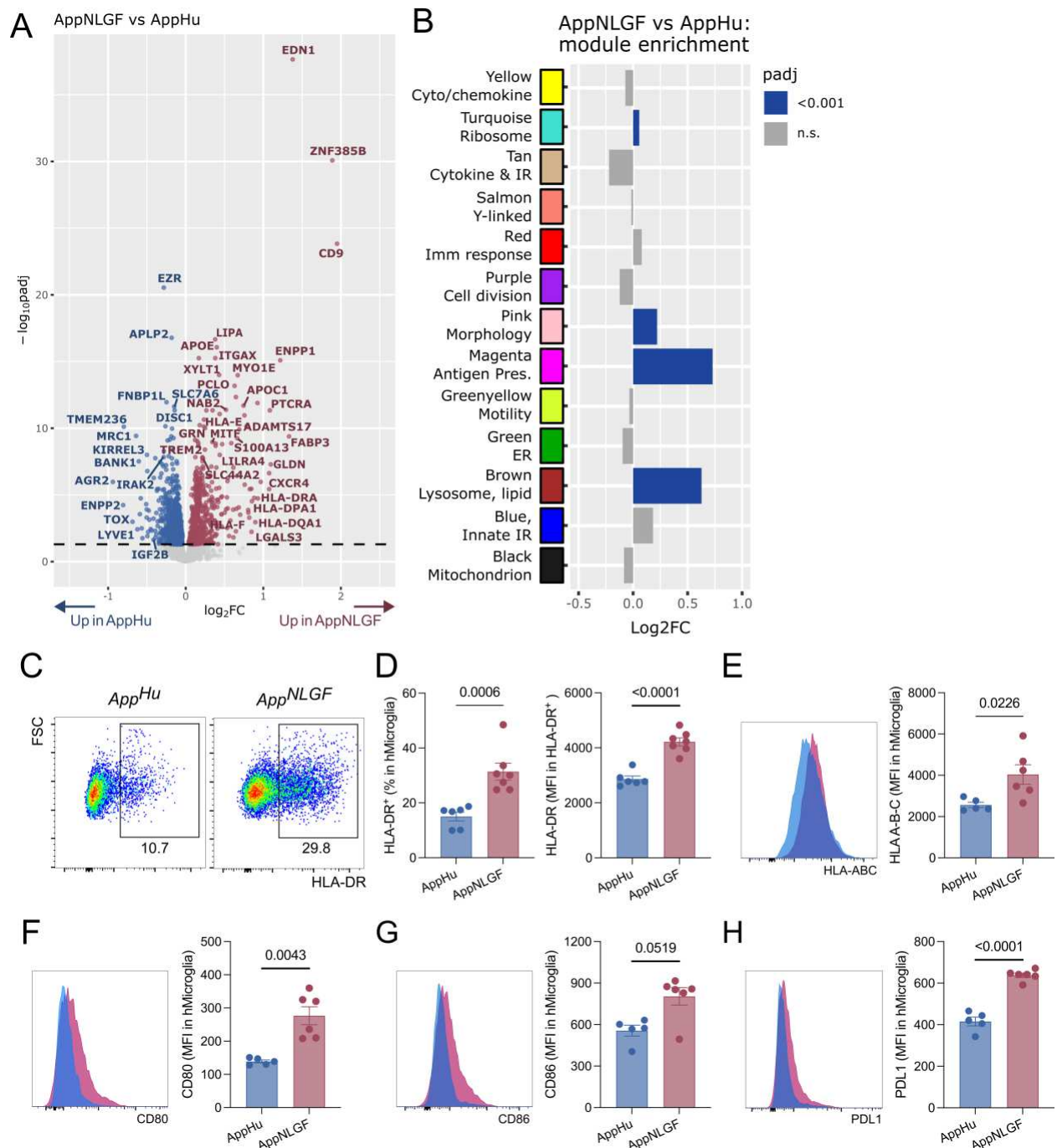


Fig. 3. The cohort-wide microglial response to amyloid includes a robust huDAM and antigen-presenting signature

(A) Volcano plot showing the average transcriptomic response of hMG to A β pathology, comparing hMG from *App^{NLGF}* (amyloid) to *App^{Hu}* (control) mice. Genes significantly upregulated (red) or downregulated (blue) in *App^{NLGF}* are indicated (Benjamini-Hochberg (BH) adjusted $p < 0.05$).

(B) Differential expression analysis on the module level, showing significant differences in module expressions between hMG extracted from *App^{NLGF}* and *App^{Hu}* mice (BH adjusted p -values as

indicated). Modules with a positive log₂(fold change) contain genes that are on average upregulated in hMG from *App^{NLGF}* mice.

(C) Gating strategy to select HLA-DR-positive H9-hMG from *App^{Hu}* and *App^{NLGF-FIRE}* mice (both on *Rag2^{-/-}*, *Il2rg^{-/-}*, *hCSF1^{KI}* background).

(D) Quantification of the percentage of MHCII^{pos} hMG (left; unpaired parametric t-test) and the mean fluorescent intensity (MFI) of MHCII^{pos} hMG (right; unpaired Mann-Whitney U test) in *App^{Hu}* (n=6) and *App^{NLGF-FIRE}* mice (n=7) demonstrating increased HLA-DR expression in amyloid plaque exposed H9-derived microglia; p-values indicated per graph.

(E-H G-J) Density plot (left) and quantification (right) of the mean fluorescence intensity (MFI) of HLA-A/B/C expression (E), CD80 (F), CD86 (G) and PDL1 (H) in H9-derived hMG from *App^{Hu}* (n=5) and *App^{NLGF-FIRE}* mice (n=6), demonstrating several elements of the immune hemi-synapse expressed on microglia exposed to amyloid plaques. Statistics: Unpaired parametric t-test (E, H); unpaired Mann-Whitney U test (D, F, G); p-values indicated per graph.

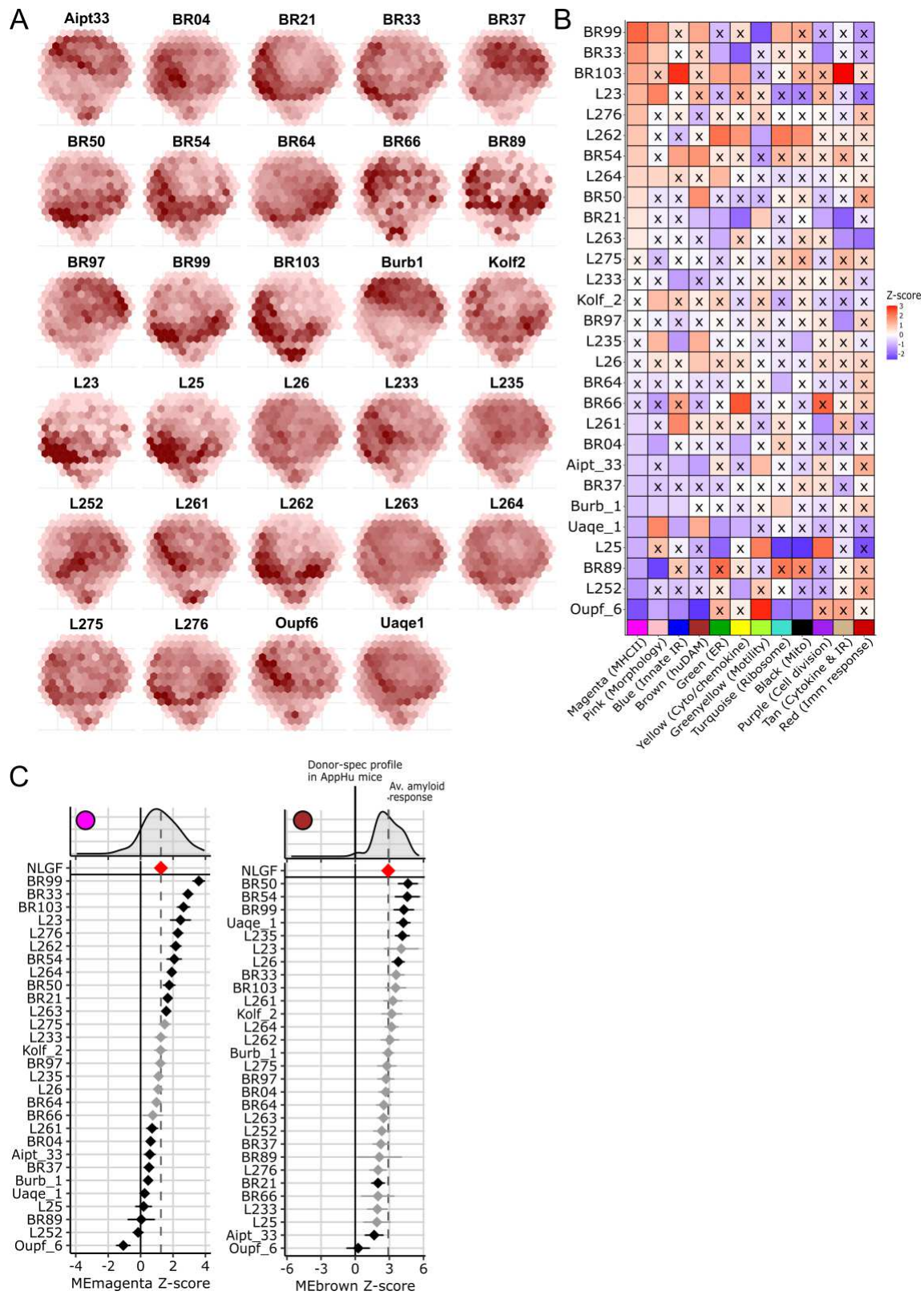


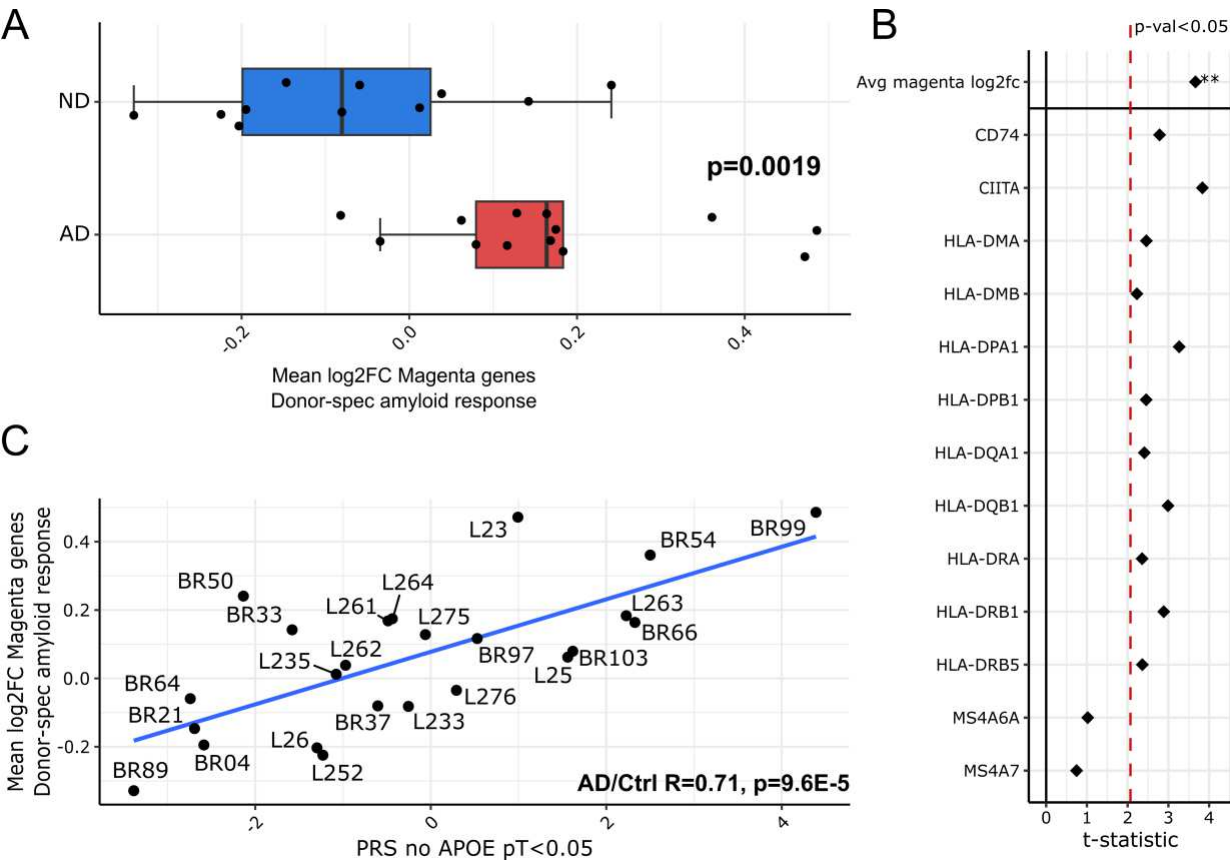
Fig. 4. MHCII activation is the dominant genetically programmed response to amyloid pathology

(A) UMAP density plots showing the distribution of hMG from each donor in *App^{NLGF}* mice, projected onto the integrated transcriptional space. Darker colors indicate where the cells of the indicated donor are distributed across transcriptional space demonstrating strong variability in the transcriptomic phenotype of the genetically diverse donors.

(B) Module differential expression analysis for each donor. Deviations from the average amyloid response derived from the cohort (Fig.3) is indicated by colour (Benjamini-Hochberg adjusted p -value < 0.05). The x-axis lists modules; the y-axis lists the individual donors. Cells marked with 'x' indicate no significant difference from the average response in Fig.3.

(C) Donor-specific z-score shifts in Magenta (MHCII; left) and Brown (huDAM; right) module expression in hMG derived from *App^{NLGF}*, which are the main components in the average amyloid response discussed in Fig.3. The average amyloid response (dashed line, indicated with a red dot in top row labeled "NLGF") and the baseline module expression in *App^{Hu}* controls (solid black vertical line at $x = 0$) are indicated, confirming the strong upregulation of the magenta (MHCII) and the brown (huDAM) response in the cohort of hMG exposed to amyloid plaques²⁰. Each diamond represents the mean module score across all cells from one donor; error bars show the 95% confidence intervals. Black diamonds indicate significant deviation from the average A β response (BH adjusted $p < 0.05$); grey diamonds are not significant. Top density plots indicate the overall distribution of module z-scores across all *App^{NLGF}* hMG.

688



689

690

691 **Fig. 5. Microglial MHCII activation in response to amyloid reflects donor AD status and**
692 **polygenic risk**

693 (A) Mean log2(fold change) of the Magenta (MHCII) module genes for the donor-specific
694 response to Aβ in AD cases (red) and non-demented controls (ND, blue). Unpaired t-test, (t(23)=
695 3.53, p=0.0019).

696 (B) Gene-wise comparison of Magenta module expression between AD cases and non-demented
697 controls in their response to Aβ pathology. Top row “Avg magenta log2fc” recapitulates the
698 module-level comparison from panel B (p=0.0019). Lower rows show individual gene-level t-
699 statistics; black diamonds to the right of the red dashed line indicate nominal significance (p <
700 0.05; uncorrected; see Methods).

701 (C) Correlation between the mean log2(fold change) of the Magenta (MHCII) module genes in the
702 donor-specific response to Aβ with the donor’s polygenic risk score for AD (PRS, excluding the
703 APOE locus; p-value threshold (pT) < 0.05), for AD cases and controls (R = 0.71, p = 9.6E-5). The
704 samples from the HipSci cohort, for whom AD status is not known, are shown in Fig.S9A (R =
705 0.71, p=0.17).

706

Table 1.

Overview of all iPSC lines used in the current study, with their official name, which cohort they are from, which village they were grafted in, their demographic details and PRS (p-value threshold <0.05) and ORS (oligogenic risk score; p-value threshold $p < 5e-08$) both excluding the APOE region. *: donors L234 and BR83 were removed from analysis due to low cell recovery.

Main_ID	Official_name	Cohort	Village	Status	APOE	Sex	PRS.noAPOE	ORS.noAPOE
BR89	AJ0089-02-SV-008	ROS/MAP	Village2	Ctrl	E2/E3	F	-3.38	-0.50
BR64	AJ0066-01-SV-007	ROS/MAP	Village3	Ctrl	E3/E4	F	-2.74	-0.06
BR21	AJ0046-01-SV-023	ROS/MAP	Village3	Ctrl	E3/E3	F	-2.69	-1.39
BR04	AJ0056-01-SV-004	ROS/MAP	Village3	Ctrl	E2/E3	F	-2.59	2.06
BR50	AJ0041-01-SV-033	ROS/MAP	Village3	Ctrl	E3/E4	F	-2.14	-1.21
Oupf6	HPSI0514i-oupf_6	HipSci	Village1	NA	E3/E3	F	-1.61	-1.23
BR33	AJ0047-01-SV-009	ROS/MAP	Village2	Ctrl	E3/E3	M	-1.58	-0.10
L26	KP-21-26	UZ Leuven	Village1-3	Ctrl	E3/E3	M	-1.30	-2.04
L252	LUMC0252iCTRL	UZ Leuven	Village1	Ctrl	E3/E3	F	-1.23	-1.55
L235	LUMC0235iCTRL	UZ Leuven	Village1	Ctrl	E3/E3	M	-1.08	1.30
L262	LUMC0262iCTRL	UZ Leuven	Village2	Ctrl	E3/E3	F	-0.97	-1.05
Kolf2	HPSI0114i-kolf_2	HipSci	Village1-3	NA	E3/E3	M	-0.78	-0.75
BR37	AJ0031-02-SV-012	ROS/MAP	Village3	Ctrl	E3/E3	F	-0.60	-0.20
L261	LUMC0261iAD	UZ Leuven	Village2	AD	E3/E3	M	-0.49	1.80
L264	LUMC0264iAD	UZ Leuven	Village2	AD	E4/E4	M	-0.44	0.19
L233	LUMC0233iAD	UZ Leuven	Village1	AD	E3/E3	M	-0.25	1.23
L275	LUMC0275iAD	UZ Leuven	Village2	AD	E4/E4	M	-0.06	-0.19
L276	LUMC0276iAD	UZ Leuven	Village2	AD	E4/E4	F	0.29	1.92
Aipt33	HPSI0513i-aipt_33	HipSci	Village1	NA	E3/E3	M	0.33	0.70
BR97	AJ0123	ROS/MAP	Village3	AD	E3/E3	M	0.53	-0.33
L23	KP-21-23	UZ Leuven	Village1	AD	E3/E3	F	0.99	1.84
L25	KP-21-25	UZ Leuven	Village1	AD	E3/E3	F	1.56	1.66
L234*	LUMC0234iAD	UZ Leuven	Village1	AD	E3/E3	F	1.60	2.07
BR103	AJ0121-01-SV-004	ROS/MAP	Village2	AD	E3/E3	F	1.62	0.94
Uaqe1	HPSI0813i-uaqe_1	HipSci	Village1	NA	E3/E3	F	1.87	-1.27
L263	LUMC0263iAD	UZ Leuven	Village2	AD	E4/E4	M	2.23	0.65
BR66	AJ0094-01-SV-010	ROS/MAP	Village3	AD	E3/E3	M	2.33	-1.09
Burb1	HPSI0714i-burb_1	HipSci	Village1-3	NA	E3/E3	M	2.33	0.55
BR54	AJ0048	ROS/MAP	Village3	AD	E4/E4	F	2.50	-1.29
BR83*	AJ0083-01-SV-011	ROS/MAP	Village3	AD	E3/E3	M	2.69	1.60
BR99	AJ0107-01-SV-013	ROS/MAP	Village2	AD	E3/E4	F	4.39	0.59

Methods

Mice

All protocols concerning animal experimentation were approved by the Ethical Committee of Laboratory Animals of KU Leuven (project no. 125/2022 & license LA1210591), following Belgian and European Union guidelines. *App*^{NL-G-F/NL-G-F} mice²⁷ carry the humanized Ab sequence which contains the Swedish (NL), Arctic (G) and Iberian (F) mutations, resulting in progressive amyloidosis and plaque formation from 2 months of age (M) and learning and memory impairments from 6M. The *App*^{Hu/Hu} mice were developed by our lab to serve as WT control, carrying the humanized A β sequence, yet without any FAD mutations²⁸. To enable hMG xenotransplantation, both strains were crossed onto the *Rag2*^{tm1.1Flv} *Csf1*^{tm1(CSF1)Flv} *Il2rg*^{tm1.1Flv}/*J* homozygous background (Jackson Laboratory, strain 017708) to obtain *Rag2*^{-/-} *Il2rg*^{-/-} *hCSF1KI***App*^{NL-G-F/NL-G-F} (henceforth named *App*^{NLGF}) and *Rag2*^{-/-} *Il2rg*^{-/-} *hCSF1KI***App*^{Hu/Hu} mice (*App*^{Hu} for short). Mice had ad libitum access to food and water and were housed in groups of 2-5 mice on a 14/10h day/night cycle at 21°C and 32% humidity. Both male and female mice were used in this study, trying to keep the ratio males to females equal across experimental conditions.

Donor selection

All participants signed informed consent and ethical approvals for genotyping and iPSC generation were granted by the Ethics Committee of the University Hospitals Leuven, Belgium (study protocol 63481).

Donors were selected from various cohorts: 1) Human iPSC Initiative (HipSci³⁴), which includes 477 healthy donors; 2) UZ Leuven Memory Clinic cohort of 200 biomarker-proven AD cases and 135 spouse controls⁶²; 3) UZ Leuven F-PACK (Flemish Prevent AD cohort KU Leuven) cohort, a community-recruited longitudinal observation cohort of 180 older adults, who undergo 2-yearly neuropsychological evaluation and ¹⁸F-flutemetamol amyloid PET imaging⁶³; 4) ROS/MAP iPSC cohort, comprising 53 deceased individuals which were both clinically and neuropathologically confirmed AD cases and controls³³. If not already available, either publicly or in-house, SNP array data was generated for each individual (Global Screening Array (GSA), Illumina).

A polygenic risk score (PRS) was calculated for each individual per cohort as previously described⁸, using the clumping and thresholding method and the PLINK genetic data analysis toolset⁶⁴ with a p-value threshold for association to AD p<0.05 on LD-clumped SNPs, where variants with R²>0.1 in a 1000-kb window were excluded and SNPs with the smallest p-value were retained. As we aimed to understand how genetic risk would impact hMG function beyond the effect of APOE, we used the PRS.noAPOE model where the PRS is calculated excluding the APOE region (chromosome 19:44.4-46.5 Mb). The scores of all cohorts were standardized against the Caucasian subset of the 1000 Genomes project, meaning that all datasets were merged, the principal components were derived and adjusted for, and the scores in different cohorts were

standardized using the mean and standard deviations from the 1000 Genomes subset. For each cohort we selected chromosomally male and female individuals (n=5 (HipSci), n=13 (UZ Leuven) and n=13 (ROS/MAP)) to span a range of PRSs yet predominantly selecting AD cases with a positive PRS and controls with a negative PRS (see also Table 1).

Human stem cells

For an overview of all iPSCs used in this study, see Table 1, which includes information regarding donor sex, APOE status, study cohort and PRS. The iPSCs from the HipSci and ROSMAP cohorts were generated as described previously^{33,34}. iPSCs from the UZ Leuven cohorts were either generated in house (L23, L25, L26) in collaboration with the KU Leuven Stem Cell Institute or at the Leiden hiPSC Centre, always starting from peripheral blood mononuclear cells (PBMCs).

For those generated in house, the PBMCs were first cultured StemPro-34 SFM Medium (Gibco, 10639-011), supplemented with penicillin-streptomycin (Gibco, 15140-122) and cytokines (SCF 100 ng/ml, Peprotech, 300-07), FLT-3 (100 ng/ml, Peprotech, 300-19), IL-3 (20ng/ml; Peprotech, 200-03) and IL-6 (20 ng/ml, Peprotech, 200-06) for 4 days to enhance expansion of erythroid lineage cells. Reprogramming was performed using 2.5×10^5 PMBCs in a 12-well plate and the CytoTune™-iPS 2.0 Sendai reprogramming kit (ThermoFisher Scientific, A16517), following the manufacturers standard protocols for feeder-free reprogramming of PBMCs. The Sendai viruses were removed the following day by centrifugating the cells (200g, 10min) and replating in fresh StemPro-34 medium with cytokines. On day 3 after reprogramming, cells were seeded in different densities (1×10^4 , 2.5×10^4 , 5×10^4) in a Matrigel coated (Corning, 354277) 6-well plate with complete StemPro34 medium without cytokines. Following daily half media changes on days 4-6, cells were transitioned to complete Essential E8 Flex Basal medium (Thermo Fisher Scientific, A28585-01; supplemented with penicillin-streptomycin 1/1000) on day 7. Colonies appeared between day15-21 after transduction, which were picked, gradually expanded and checked for remaining Sendai virus using qPCR (TaqMan® iPSC Sendai Detection Kit, A13640) from P9 onwards. Quality checks included mycoplasma, HIV & hepatitis B/C testing, testing for pluripotency and stemness using immunocytochemistry (OCT4 (Abcam, ab19857, 1/200), Tra1-81 (Milipore, mab4381, 1/200), SOX2 (Abcam, ab92494, 1/200), Nanog (R&D systems, AF1997, 1/200), OTX2 (Millipore, AB9566, 1/500), Ki67 (BD Pharmingen, 556003, 1/200) and the TaqMan HPSC scorecard method (Life Technologies – A15872) after trilineage differentiation via embryoid bodies formation, as per manufacturers protocol, and karyotyping using aCGH array (Cytosure Syndrome Plus 180K array).

For those iPSCs generated by the Leiden hiPSC Centre, the PBMCs were reprogrammed using episomal polycistronic lentiviral vectors with or without p53 knockdown⁶⁵ after which they were cultured on Matrigel or Vitronectin XF (STEMCELL Technologies, cat. #100-0763) and mTESR™ Plus (Stemcell Technologies, cat. #100-0276). Stem cells were quality checked at Leiden hiPSC Centre for stemness and pluripotency using flow cytometry (NANOG, SSEA4, OCT4) and immunocytochemistry (PAX6, FABP7, Nestin, GATA4, FOXA2, EOMES, Vimentin, CDX2 and Brachyury). Additional quality checks included mycoplasma testing, HIV, hepatitis B/C

testing and karyotyping using aCGH array. All stem cells were transitioned to Matrigel coating and complete Essential E8 Flex Basal medium upon arrival, expanded and banked in liquid nitrogen until further use.

Stem-cell-derived microglia generation and xenotransplantation

Human MG precursor cells were generated using the MIGRATE protocol as previously published²⁹. For the xenotransplantations where 1 line was grafted per mouse (from here on referred to in Materials & Methods as the ‘single grafts’), we performed the MIGRATE protocol in batches of 1-4 donor lines and always including differentiation of the Kolf2 line as an internal control for batch variation. For the hMG village experiments, precursor cells from 12-13 different donors were generated for each line separately, harvested and counted on day 18 and pooled in equal ratios where numbers permitted. Every hMG village experiment contained cells from donors Kolf2, Burb1 and L26 to facilitate data integration across hMG villages. In total we generated 3 separate hMG villages that were differentiated independently. hMG precursors were xenografted in pups of postnatal day 4, who had received BLZ945 (200 mg/kg;) i.p. in the 2 days prior to xenotransplantation, injecting 250.000 cells in 1µl per hemisphere, aiming for +1mm ML and -1mm AP from Bregma. The same hMG village was grafted simultaneously in both *App^{Hu}* and *App^{NLGF}* mice. For the single graft experiments we only grafted *App^{NLGF}* mice, grafting both the donor lines within that batch as well as the Kolf2 line alongside (in separate mice) to control for batch effects. In total, hMG from n=16 donor lines were grafted for the single graft experiments and villages n=3 were grafted in n=33 mice.

Human microglia isolation from mouse brain

At 6M of age, mice were sacrificed using an overdose of sodium pentobarbital followed by transcardial perfusion with 20 ml of ice-cold DPBS (Gibco, Cat #. 14190-144) containing 5 units of heparin (Heparin Leo 25000 IU/5 ml, Leo Pharmaceuticals). Each of the three hMG village experiments were transplanted into 4-6 mice per strain (*App^{Hu}* or *App^{NLGF}*). For the single graft experiments we sacrificed n=2-6 mice (*App^{NLGF}* only) total per donor. For Kolf2 we sacrificed a total of n=24 mice, sacrificing n=2 Kolf2 mice per injection batch of single donor transplantations (see previous section) to control for batch variation. The right hemisphere (omitting cerebellum) was placed in ice-cold FACS buffer (1x DPB, 2% FBS, 2mM EDTA) including 5µM actinomycin D (ActD; Sigma, A1410-5MG), to prevent artificial microglial activation due to the isolation procedure, as previously described⁶⁶. The samples were mechanically and enzymatically dissociated using the Miltenyi Neural Tissue Dissociation Kit P (Miltenyi, 130-092-628) supplemented with 5 µM ActD, filtered using a 70 µM strainer (Greiner Bio-One, 542070), washed with 15 ml of ice-cold FACS buffer + 5µM ActD and centrifuged (300g, 15min, 4°C). From here on, ActD was omitted from the buffers. Next, a sample clean-up was performed by resuspending the pellet in a 30% isotonic Percoll gradient (Fischer Scientific, cat. #11500744), centrifuging (300g, 15min, 4°C) and removing the layers of myelin and cellular debris that accumulate on top

of the sample. The Fc receptors on the cells were blocked in FcR blocking solution (1/10, Miltenyi, 130-092-575) in ice-cold FACS buffer for 10 min at 4°C. After a 5ml FACS buffer wash and centrifugation (300g, 5 min, 4°C), pelleted cells were incubated for 30 min at 4°C in FACS buffer with the viability dye eFluor780 (1/2000; Thermo Fisher Scientific, 65-0865-14) and the following antibodies: PE pan CD11b (1/50; Miltenyi, 130-113-806), APC human CD45 (1/50; BD Biosciences, 555485), BV421 mouse CD45 (1/500; BD Biosciences, 563890) and Total-Seq A cell hashing antibodies (1/500, Biolegend). Following incubation, cells were washed, pelleted through centrifugation (300g, 5 min, 4°C), resuspended in 500 µl ice-cold FACS buffer and passed through a 35 µm strainer, before sorting on the MACSQuant Tyto. Human microglia (hMG) were considered positive for CD11b and hCD45 and per mouse 40.000 hMG were collected and diluted to 1000 cells/µl in FACS buffer.

Single cell RNA library preparation and sequencing

As cells were hashed using Total-Seq A cell hashing antibodies³⁶, cells from different mice could be pooled. For the hMG Village experiments cells from the different mice (n=5-6 per sacrifice date) were pooled in equal numbers. For the single graft experiments, cells from different mice were combined to enable retrieval of roughly equal number of cells from 1 donor across variable mouse numbers (n=2-6). Given the high number of Kolf2-grafted mice (n=24), we only loaded 2000 cells per Kolf2-grafted mouse. The LUNA dual fluorescence cell counter (Logos Biosystems, Villeneuve d'Ascq, France) was used to verify cell density and viability of the samples. Pooled cells were loaded onto a custom in-house HyDrop microfluidics droplet generator⁶⁷. In short, the desired cell volume, the single cell barcoding mix and the barcoding gel beads were loaded separately into mineral oil-backed syringe tips and the droplet generating partition oil was also loaded into a separate syringe. Next, the syringes were loaded onto independent syringe pumps on the ONYX droplet platform (Atrandi Biosciences, Vilnius, Lithuania) and droplets were generated with the in-house developed HyDrop microfluidics chips⁶⁷. The ability of fine-tuning the liquid flowrates with custom HyDrop microfluidics setup and the ability to observe real-time droplet encapsulation allows to get higher cell encapsulation efficiency and lower failures resulting from encapsulation anomalies.

Single cell RNA-seq libraries were subsequently prepared using 10X Genomics Chromium Single Cell 3' Kit, v3.1 NextGEM chemistry (CG000204 Rev D, 10X Genomics, Pleasanton, CA, USA) following manufacturer's instruction. Hashtag oligo (HTO) libraries were prepared in parallel also following manufacturer's instruction (BioLegend, Total-Seq A Antibodies and Cell Hashing with 10x Single Cell 3' Reagent Kit V3.1 Protocol), using 16 cycles for the index PCR. The quality of the libraries was assessed using the Qubit (Thermo Fisher Scientific, Schwerte, Germany) and Bioanalyzer (Agilent Technologies, Santa Clara, USA). In total we created 24 libraries for the hMG Village experiment, encompassing 33 mice in total, and 33 libraries for the single grafts, encompassing 99 mice in total. Libraries were sequenced aiming for 90% mRNA and 10% HTO libraries (55,000 total reads per cell) on a NovaSeq6000 (Illumina) with the read lengths recommended by 10X Genomics.

Data processing, quality control and dimensionality reduction

The sequencing reads were aligned to a combined murine mm10 and human GRCh38 genome. Analysis of single cell sequencing data was conducted in R version 4.3.3 and Seurat V5⁶⁸. HTO computational deconvolution was done using Seurat. For the village grafts additional donor deconvolution was performed based on the unique SNP profile of each donor, using SoupOrCell³⁷. This method assigns donor origins by comparing known SNPs between donors to those detected in the sequencing reads of individual cells. These cells are then clustered based on their SNP profile. Next, a mixture model is applied to assign a cell to the donor of origin, to identify if it is a genetically heterogeneous doublet, or conclude that insufficient information is available for a confident donor assignment. To validate the SNP-based deconvolution approach, we used data from the single grafted lines, as here individual mice are grafted with a single donor but also received a Total-Seq A cell hashtag, allowing us to compare the SNP-based deconvolution to antibody-based deconvolution. The overlap was on average 99% (data not shown).

Next, we removed low-quality cells that could not be confidently assigned to a single donor, that had too few Total-Seq A cell hashtag reads to be confidently mapped back to their murine host, and those with more than 10% murine reads. Additionally, we used miQC⁶⁹ to generate data-driven mitochondria percentage cut-offs for each sequencing library and removed any offending cells. For the village data, the proportions of cells retrieved from each donor were variable, but highly consistent across mice, and for those lines included in each village, quite consistent across villages (Fig.1B, Fig.S1A). Donors with fewer than 500 cells in the village data were dropped from the analysis (L234, BR89). For the single graft data, we dropped the cells from donor L262 as the data was derived from only 1 mouse.

Next, we removed any cells with <200 features and >5000 features as they are indicative of uninformative cells and doublets, respectively. Finally, we performed an initial clustering of the cells and removed clusters of proliferating cells and contaminating border associated macrophages (Fig.S1B). After all quality control steps, we retained n=235,106 microglia from 29 human donors, grafted across 33 mice.

The raw RNA counts from each sequencing libraries were normalized using SCTransform (SCT)⁷⁰. Dimensional reduction was performed on the SCT residuals using Principal Component Analysis (PCA) and optimal PCA retention was determined by elbow plot analysis, retaining thirteen principal components (PCs). Sequencing batch effects were integrated out using Harmony⁷¹. Using the integrated space we constructed 2D UMAPs. These UMAPs were binned to produce the hexbin module score (see below) and donor density plots. For module scores (see below), each hexagon represents the mean module expression of all cells within that area of the UMAP. Patient density plots depict the total number of microglia in each hexagon relative to the total number of cells recovered from each donor.

WGCNA

Modules of co-expressing genes were identified by Weighted Gene Co-expression Network Analysis (WGCNA⁷²). Similarity scores were calculated using Pearson's correlations among the 5,000 most variable genes as identified by SCT. The default WGCNA approach was used to determine the power that best approximated a scale-free network topology, with a minimum cluster size of five and a default splitting parameter of two. This analysis yielded thirteen gene co-expression modules, which were functionally annotated using overrepresentation analysis (Fisher's exact test with Benjamini-Hochberg correction for multiple testing) by comparing their member genes to known microglial cluster markers as previously published²⁰.

Differential expression testing

For the microglia village dataset, we analyzed differences in gene expression using a generalized linear mixed effects model (GLMM) with a negative binomial distribution, which allowed us to assess the impact of both biological and technical variables. Our experimental design included three reference donors in each sequencing batch and grafted village (L26, Kolf2, Burb1) to ensure the model remained identifiable. The model was applied directly to the unaltered count matrix. Sequencing batch and host sex were included as fixed effects within the statistical framework, while an offset was taken to control for sequencing depth. To account for the non-independence of cells from the same host animal, we included a random intercept for mouse host, thereby controlling for pseudo-replication and reducing false-positive discovery rates. We used glmmTMB⁷³ to efficiently fit the model across the 10,000 most variable genes as determined by SCT. Genes that failed to converge or produced non-positive-definite Hessian matrices, mostly due to low coverage, were excluded from downstream analysis. Model in pseudo code:

$$\begin{aligned} y &\sim \text{negbin}(\lambda, \phi) \\ \log(\lambda) &\sim \text{donor} + \text{mouse strain} + \text{donor} * \text{mouse strain} + (\text{mouse id}) + \text{mouse sex} + \text{batch} + \log \text{depth} \\ \text{host id} &\sim \mathcal{N}(\mu, \sigma) \end{aligned}$$

From this model we extracted four contrasts of interest: 1) the donor effect: genes differentially expressed (DE) in one donor compared to the grand mean of all donors (i.e. the mean of all donor averages) only in *App^{Hu}* mice reflecting basic genetic differences; 2) the average amyloid response: equivalent to the mouse host strain effect identifying DE genes between all hMG from *App^{NLGF}* mice versus *App^{Hu}* mice; 3) the personalized amyloid response: equivalent to the donor*host strain interaction effect, quantifying how each donor deviates from the average amyloid response, that is, genes significantly up- or downregulated in a specific donor relative to the cohort's mean amyloid response; 4) the donor sex effect: the DE between all male and female donors. All contrasts were derived from the fitted model using the emmeans package in R⁷⁴. The resulting parameter estimates were post-hoc corrected for donor sex using a simple linear model (given that sex is collinear with donor, it could not be fitted in the model), and adjusted p-values were derived from the corrected estimates and Benjamini-Hochberg FDR corrected to account for the inflation in false positives due to multiple testing.

Many differential expression methods for single cell sequencing data are known to have inflated false positive rates due to sample size inflation³⁸. To evaluate the robustness of the model in controlling type I errors, we conducted a randomization experiment. Donor information was randomized while maintaining the same proportions within sequencing batches and mouse hosts, ensuring no true biological effects existed. We then performed a differential expression analysis (using the grand mean across all donors as a reference) assessing the impact of donor, while retaining potential confounding assignments like sequencing batch. Our model produced maximally 1 false-positive per comparison, indicating strong control over false positive rates (Fig.S1C). We compared this with the performance of the often-used Wilcoxon rank sum test (using donor Uaqel as a reference for the DE). This DE model produced thousands of false positives across donors, especially when comparing across sequencing batches (Fig.S1C). The Wilcoxon method assumes, in contrast to the GLMM method we applied here, statistical independence across cells. From a biological (baseline inter-mouse variability in gene expression) or technical perspective (e.g. batch effects) this is not tenable, and when Wilcoxon is used for the analysis of a large population (>250,000 cells) this results in many false positive changes as illustrated here (Fig.S1C).

For the module-level differential expression analysis, we employed a linear mixed effects model (LMEM), using continuous module scores (eigengenes) derived from WGCNA. The same covariates and adjustments for sex differences were incorporated, maintaining consistency across analyses.

Gene Set Enrichment Analysis

To identify enriched pathways in our DE results and WGCNA modules, we employed the Fast Gene Set Enrichment Analysis (FGSEA) package⁷⁵. Where possible we used the signed $\log_{10}(\text{pvalue})$ as the ranking metric to identify the enrichment of groups of genes in our differential expression based on their annotation in the Gene Ontology (Biological Processes) and KEGG databases. To examine the function of WGCNA modules, we compared the genes present in each module against these same databases using overrepresentation analysis (Fisher's exact test).

Comparing single grafts to village grafts

The single graft experiment contained hMG from 16 donors that were xenotransplanted individually into *App^{NLGF}* mice (1 donor per mouse, 2-6 mice per donor). Cells from donor L262 were dropped (see "Data processing, quality control and dimensionality reduction"). DE analysis was performed on this dataset (n=299,368 good quality hMG) comparing the transcriptome of each donor to that of L235, considering we had sufficient cells of L235 in both the single graft experiment (n=22,498) as well as in the village experiment (n=19,849). To make the village data comparable, we subsetting the village data to include only cells from the *App^{NLGF}* mice, and performed the DE analysis comparing each donor to L235. Next, the gene expression profile of each donor between the two datasets was compared using Pearson's correlation.

Association analysis with the Magenta personalized amyloid response

The association between the Magenta donor-specific amyloid response and AD status was assessed using a Welch Two Sample t-test comparing the mean log2 fold change (LFC) of all Magenta genes (sex-corrected) between AD patients and non-demented controls. Five donors with unknown AD status (HipSci cohort) were excluded from the analysis. We further investigated whether the mean LFC of the genes in the Magenta module in the personalized amyloid response (sex-corrected) per donors were correlated with their polygenic risk scores using a Pearson's correlation.

MHCII haplotyping

DNA extraction was performed on iPSC pellets of the same donor and clone as used in the hMG village experiments, using the PCI method. In short, cell pellets were lysed in cell lysis buffer (10mM Tris (pH 8), 100mM NaCl, 10mM EDTA (pH 8), 10% SDS) containing proteinase K (20 mg/ml) for 1h at 56°C with agitation (1400rpm). After adding phenol:chlorophorm:isoamyl alcohol (25:24:1, pH 7.8-8.2; Fluka Biochemika, cat# RA14615), vortexing for 1min and spinning at max speed for 5min, the aqueous phase was removed into a fresh tube. Additional retrieval was performed by adding elution buffer (10mM TrisHCl pH 8.5) to the cell lysate tube, vortexing and spinning at max speed, and removing the additional aqueous phase, which was merged with the initial aqueous phase. Next, equal volumes of chloroform:isoamyl alcohol (Sigma, cat# 0549-1PT) was added to the aqueous phase, samples were vortexed and spun at max speed for 5 min, after which the aqueous phase was again transferred to a clean tube. Ethanol precipitation was achieved by adding 0.75M NH4OAc (final concentration) and 20 µg of GlycoBlue (Invitrogen, cat# AM9516) to the sample, mixing and subsequently adding 2.5 volumes of 100% ethanol, mixing and incubating at -80°C for 30min. After a 20 min spin at full speed at 4°C, the supernatant was decanted, the samples were washed 2x in 80% ethanol and dried overnight at RT. Samples were resuspended in 60°C elution buffer and stored at -20°C until further handling. Genotyping of all samples to G-level resolution of the MHCII locus (exon 2 and 3 for HLA-A, -B, -C and exon 2 for HLA-DRB1₃₄₅, -DQA1, -DQB1, DPA1, -DPB1) was performed using the HiSeq sequencing system (Illumina Inc., San Diego, USA). The IMGT/HLA allele database version 3.55.0 was used for results reporting.

Immunofluorescence on xenograft mouse brain

The left brain-hemisphere of the same mice used for the single graft hMG isolation were immersion fixed overnight in 4% PFA and cut in 40 µm thick sections on the vibratome. Sections were permeabilized (15min, RT with shaking in PBS 0.2% Triton X100 (PBST)), stained with X34 staining solution (10µM final concentration in 40% ethanol in PBS (vol/vol) and 20mM NaOH), washed 3x 2min with 40% ethanol in PBS (vol/vol) and 2x 5min with PBST, blocked in blocking solution (5% donkey serum in PBST) for 1h at RT and incubated overnight with primary antibody at 4°C (anti-human HLA-DR/DQ/DP, 1/200, cat.# ab7856, Abcam) in blocking solution.

After washing (3x 5min PBS), the sections underwent 2h of secondary antibody incubation at RT (donkey anti-mouse Alexa647 (1/500, cat# A31571, Invitrogen) followed by washes (3x 5min PBS), 1h blocking and another round of primary antibody incubation overnight at 4°C (anti-human CD9 biotin (1/100, cat.# 312112, Biolegend), anti-human P2RY12 (1/1000, cat.# HPA014518, Atlas Antibodies)), followed by washes (3x 5min PBS), incubation with the secondary antibodies for 2h at RT (donkey anti-rabbit 594 (1/500, cat.# A21207, Invitrogen), streptavidin 488 (1/500, cat.# S32354, Thermo Scientific)) and mounting in Glycergel (C0563, Dako). Confocal images were obtained using the Nikon AX inverted microscope driven by NIS software (v5.42.06), exciting at 405, 488, 561 and 640nm and compiled using Fiji/Image J software.

Quantification of MHCII in human brain samples

Human post-mortem brain

Brain tissue from 15 human autopsy cases was included for this study (Table S8). The clinical Dementia Rating (CDR) score, reflecting the stage of cognitive and functional impairment, was retrospectively assessed based on clinical files. Patients met the criteria for (a) symptomatic Alzheimer's disease (AD), defined by clinical signs of dementia and intermediate to high levels of AD neuropathological changes (as described in ⁷⁶), or (b) the absence of dementia symptoms along with AD neuropathological changes. Autopsies were conducted at university or municipal hospitals in Belgium (Leuven) and Germany (Bonn and Offenbach) in compliance with local laws. Informed consent was obtained in accordance with local legislation. Ethical approval for the recruitment protocols and collection of human brain tissue was granted by the ethical committees of UZ Leuven (Belgium; S59292, S52791) and the University of Ulm (Germany, 58/08). The use of the brain sample for the experiments reported here was approved by the ethical committee of UZ-Leuven (S-59295, S-65147). The right hemisphere of the cerebrum, cerebellum, and the right half of the brainstem were dissected for gross neuropathological examination and stored at -80°C. The left cerebral hemisphere, left cerebellar hemisphere, and left half of the brainstem were fixed in 4% aqueous or phosphate-buffered formaldehyde for 2 to 4 weeks, after which they underwent dissection and gross neuropathological assessment. Tissue samples from the frontal cortex (Brodmann area (BA) 6) were collected during autopsy.

Chromogen immunohistochemistry on human post-mortem brain

Immunohistochemistry was performed on tissue samples from the frontal cortex. Information and dilutions of the primary antibodies can be found in Table S9. Paraffin-embedded tissue sections were deparaffinized and subjected to heat-induced epitope retrieval (pH 6.1) followed by incubation with Envision Flex Peroxidase-Blocking Reagent (Dako) for 5 minutes. Tissue sections were then incubated with primary antibodies overnight, followed by incubation with horseradish peroxidase (HRP)-conjugated secondary antibodies (EnVision+ System HRP-Labelled Polymer Anti-mouse, Dako) or biotinylated secondary antibody (REAL Detection System Alkaline Phosphatase/RED Rabbit/Mouse, Dako) and subsequently alkaline phosphatase-conjugated streptavidin for 30 minutes. MHCII antigen-antibody complexes were visualized using 3,3'-

Diaminobenzidine tetrahydrochloride hydrate (DAB; Liquid DAB+ Substrate Chromogen System, Dako) or Vector SG Substrate Peroxidase (Vector SG Substrate Kit, Peroxidase (HRP), Vector Laboratories), Fast-Red (REAL Detection System Alkaline Phosphatase/RED Rabbit/Mouse, Dako) was used for the visualization of 4G8 immunoreactivity. CD3 antigen-antibody complexes were visualized using DAB. Incubation with Envision Flex Peroxidase-Blocking Reagent (Dako) was repeated before incubation with subsequent primary antibodies overnight. Slides were counterstained with hematoxylin, dehydrated using an autostainer (Leica Biosystems), and mounted with an automated cover-slipper and Leica CV mount (Leica Biosystems). Microscopic images were captured using a Leica DFC7000 T camera (Leica Microsystems, Wetzlar, Germany) mounted on a Leica DM2000 LED light microscope (Leica Microsystems). Image processing was performed using ImageJ software (National Institutes of Health, Bethesda, USA) and Inkscape (<https://inkscape.org/>).

Quantification of microglia with MHC class II upregulation

Microglial cells immunoreactive to MHC class II were quantified in layer V/VI of the frontal cortex, after immunohistochemical staining for HLA-DR/DP/DQ using Vector Substrate Peroxidase, 4G8 using Fast-Red and CD3 using DAB. Only regions without infarcts, microinfarcts or bleedings were considered for quantification. Microglial cells were morphologically distinguished based on cell type-specific cytological features as described by García-Cabezas *et al.*⁷⁷. For each AD case, twenty regions of interest were determined with a diameter of 100 µm around Aβ plaques, and twenty regions of interest of 100 µm in areas without Aβ plaques. For non-AD cases, twenty regions of interest with a diameter of 100 µm were determined in areas without Aβ plaques, containing glial cells. Images were taken using a 40x objective, mounted on a Leica DM2000 LED microscope (Leica Microsystems). Cell counts were performed manually using ImageJ software.

hMG & T cell co-culture

Mice

H9 embryonic stem cells (WA09, WiCell, Madison, WI, USA) were differentiated into human microglia precursor cells and xenotransplanted as described above (“Stem-cell-derived microglia generation and xenotransplantation”) into *App^{Hu}* mice and *Rag2^{tm1.1Flv} ; Csf1^{tm1(CSF1)Flv}; Il2rg^{tm1.1Flv/J}; App^{tm3.1Tcs}; Csf1^{Rem1Bdes}; App^{NL-G-F/NL-G-F}* (¹ from here on named *App^{NLGF-FIRE}*), which carry the fms-intronic regulatory sequence (FIRE sequence) deletion in intron 2 of the mouse *Csf1R* gene, which genetically depletes all mouse microglia. At 5 (*App^{Hu}*) and 8 months of age (*App^{NLGF-FIRE}*) mice were sacrificed through an overdose of pentobarbital and intracardiac perfusion with PBS-Heparin and the hMG were isolated from brain as described above (“Human microglia isolation from mouse brain”). hMG were stained for viability with eF780 dye (Thermo Fischer, cat. #65-0865-18) and Alexa Fluor 488 anti-human CD45 (cat. #304017, Biolegend); PE/Dazzle 594 anti-human CD274 (PD-L1; cat. #329732, Biolegend); PeCy7 anti-human HLA-DR (cat. #307616, Biolegend); Brilliant Violet 421 anti-human CD80 (cat. #305221, Biolegend);

PE anti-human CD86 (cat. #305438, Biolegend); APC anti-human HLA A/B/C (cat. #307616, Biolegend). Precision Counting beads (cat. #424902, Biolegend) were added before acquisition. Samples were acquired and counted on a BD FACSymphony A1 (BD Biosciences) and analyzed using FlowJo (BD Biosciences).

Peripheral blood mononuclear cell isolation

Buffy coat samples from healthy donors were obtained from the Red Cross Donor Center Mechelen, Belgium and approved by the Ethics Committee of the University Hospitals Leuven, Belgium (study protocol S68611). Peripheral blood mononuclear cells were obtained by Ficoll density centrifugation (Axis-Shield, 1114545) and washed in PBS containing 1 mM EDTA. The ring at the interface was collected, washed with PBS, counted and frozen at 50×10^6 cells/mL in FCS 10% DMSO at -150°C .

In vitro human T cell assay

For mixed leukocyte reaction (MLR) experiments, human T cells were isolated from frozen PBMCs using the Naive Pan T Cell Isolation Kit (Miltenyi, cat. #130-097-095). T cells were plated on top of hMG at a ratio 5:1.

For proliferation and activation assessment, isolated Naive Pan T cells were labeled with $5\mu\text{M}$ Cell Trace Violet (CTV; Thermo Fischer, cat. #C34557) for 15 min at 37°C and cultured with hMG. At day 5, cells were collected and stained for viability with eF780 dye (Thermo Fischer, Cat# 65-0865-18) and extracellular markers APC anti-human CD45RO (cat. #304210, Biolegend); FITC anti-human CD8 (cat. #344704, Biolegend); PE/Cy7 anti-human CD45RA (cat. #304126, Biolegend); PerCP/Cy5.5 anti-human CD4 (cat. #300530, Biolegend) in PBS, 3% FCS, 2mM EDTA. Samples were acquired on a BD FACSymphony A1 (BD Biosciences) and analyzed using FlowJo (BD Biosciences).

For the study of cytokines produced by T cells, CytoStim (Miltenyi, cat. #130-092-172) was added during all the coculture, according to manufacturer's instructions. At day 5, cells were collected and restimulated for 5h at 37°C with Cell Stimulation Cocktail (cat. #00-4970-03, Thermo Fischer) in presence of Brefeldin A (cat. # 420601, Biolegend) and Monensin (Thermo Fischer, Cat# 00-4505-51). Cells were stained for viability with eF780 dye and extracellular markers APC anti-human CD45RO (cat. #304210, Biolegend); FITC anti-human CD8 (cat. #344704, Biolegend); PerCP/Cy5.5 anti-human CD4 (cat. #300530, Biolegend) in PBS 3%FCS 2mM EDTA. For intracellular staining the cells were fixed and permeabilized using Transcription Factor Staining Buffer Set (cat. #00-5523-00, Thermo Fischer) according to manufacturer's instructions. The cells stained with PE anti-human IFN γ (cat. #502510; Biolegend); BV421 anti-human IL-2 (cat. #500328, Biolegend); Brilliant Ultra Violet 737 TNF α (cat. #367-7349-42, Thermo Fischer)). Samples were acquired using BD LSRFortessa X20 (BD Biosciences) and analyzed using FlowJo (BD Biosciences).

1147

1148 **Statistics**

1149 Depending on the normality of the residuals and the equality of variances of the data, unpaired
1150 parametric t-test with or without Welch's correction or Mann-Whitney U tests were applied with
1151 Dunn's post-hoc testing where applicable, as indicated in the respective figure legends, with
1152 $\alpha=0.05$.

1153

1154

Supplementary Text 1 (ST1)

Sex differences in microglial biology have been reported in both mice and humans, yet their transcriptional impact in the context of natural human genetic diversity remains incompletely understood. To examine the effect of donor sex on human microglia function, we compared single-cell transcriptomes from male and female donors across all experimental conditions. We identified 1,735 significantly upregulated genes and 4,201 downregulated genes in hMG from female compared to male donors (see Fig.1C & Table S1). Among the most enriched transcripts in female microglia were multiple X-linked genes (*XACT*, *PRKX*, *ZFX*, *JPX*), as well as MHC I and MHC II genes (e.g *HLA-A*, *HLA-B*, *HLA-DRA*, *-DRB1*, *-DQB1*, *CD74*). Additional upregulated genes included members of the AD-associated MS4A locus (*MS4A6A*, *MS4A4A*, *MS4A7*, *MS4A6E*), cell-surface receptors (*CLEC17A*, *TLR7*, *TLR2*, *TNFRSF11A*), cytokines and chemokines (*TNFSF18*, *CXCL10*, *CCL20*, *IL1A*) and interferon signaling genes (*IFI27*, *IFI22L*, *ISG15*, *MX1*; see Table S1).

Gene ontology and KEGG pathway analyses confirmed overrepresentation of immune-related pathways in female microglia including antigen processing and presentation, cytokine signaling, and cell adhesion (Fig.S2A, Table S2). Conversely, male microglia showed increased expression of several Y-linked genes (*UTY*, *NLGN4Y*, *USP9Y*, *TTY14*), as well as genes encoding glycoproteins (*TNR*, *GP7*), transcriptional regulation (*TOX*, *ZNF556*) and cell adhesion (*CADM2*, *CDH10*, *PCDHGA10*). These findings suggest that human donor sex influences core immunological features of microglial identity—including MHC expression and cytokine responsiveness—and highlight the importance of accounting for sex as a biological variable in genetically diverse human microglia models.

Supplementary figure legends

Fig. S1. Quality checks and statistical models

(A) Proportion of cells per donor (as a percentage) across all mice within a single village. Note that the top 3 lines were included in all three hMG villages.

(B) Initial UMAP highlights the hMG (black) that were kept and the CNS-associated macrophages (CAMs; red) and proliferating cells (blue) that were removed from the analysis.

(C) Comparison of statistical methods for differential expression analysis on randomized versus real data (see Methods). Left panel (Green): generalized linear mixed effects model (GLMM) applied to true donor-labeled data across both *App^{Hu}* and *App^{NLGF}* mice reveals substantial donor-specific variation (using the grand mean across all donors (i.e. the average of donor averages) as a reference for the DE analysis). Middle panel: GLMM applied to randomized donor assignments, preserving batch and host proportions (using the grand mean across all donors as a reference for the DE analysis). As expected, DEG counts approach zero. Right panel (pink): the widely-used Wilcoxon test applied to randomized donor assignments (using Uae1 as a reference for the DE analysis). Despite no true biological signal, this method yields spurious DEG calls in the current experiment (also in a sequencing batch dependent manner), illustrating that without mixed-effect modeling one would inflate false positive rates.

(D) Correlation analysis comparing the anchor lines across the three different villages and with its mean across all 3 villages. Correlations for Kolf2 are low for Village 1, due to low cell number (n=310), as demonstrated by (E).

(E) Subsampling of donor BR64 to series of ~300 cells demonstrates that drops in correlation across different subsamples are predominantly due to low cell number, not due to intervillage variation, given that BR64 was only grafted within village #3.

(F) Overview of the statistical comparisons that were conducted in this study, showcasing a subset of donors and HLA-DBQ1 expression as example (*App^{Hu}*: SCT normalized expression data; *App^{NLGF}*: expression values extrapolated from the DE model). “Donor effect” (left panel): assesses to what extent HLA-DQB1 expression of a single donor only in *App^{Hu}* mice deviates significantly (depicted with the double-sided arrows) from the average HLA-DQB1 expression across all donors (dashed vertical line; representing the average of donor means). “Strain effect” (middle panel): assesses HLA-DQB1 expression in all donor cells derived from *App^{Hu}* mice and compares that to all donor cells derived from *App^{NLGF}* mice. “Strain*Donor effect” (right panel): assesses whether specific donors deviate significantly (depicted with the double-sided arrows) from the main strain effect (dashed vertical line, derived from left panel). The same statistical concepts are applied on WGCNA module data.

Fig. S2. Major transcriptional differences between male and female hMG

(A) Gene set enrichment analysis of sex-associated DEGs using GO Biological Process (GOBP) and KEGG pathway terms. Selected terms with Benjamini-Hochberg-adjusted $p < 0.05$ are highlighted. See also Table S2.

Fig. S3. Correlation structure among gene expression modules across all hMG

(A) Pearson correlation matrix of WGCNA module scores averaged across all 235,106 hMG. Nearly all pairwise correlations are statistically significant after BH correction due to the high numbers. Modules associated with immune activation (e.g., Magenta, Brown, Pink) show positive correlations, while others such as Greenyellow (motility) exhibit anti-correlation with these immune modules, reflecting structured and combinatorial gene program usage across microglia. (B) Network representation of the correlation structure depicted in A, showing only positive correlations. Size of the nodes represent the strength of the correlation. The Greenyellow module seems to branch out into 2 separate responses, one involving cytokine release and innate immunity responses, and the other involving ribosomal and mitochondrial changes together with huDAM and MHCII responses.

Fig. S4. Donor-specific profile of WGCNA module induction in the control brain

Z-score shifts in module expression across all 29 donors in *App^{Hu}* mice. Each panel represents one WGCNA module. The x-axis denotes z-score shift in module expression relative to the average hMG baseline across all donors (solid black line at $x = 0$), with 1 unit corresponding to one standard deviation. For reference, the average *App^{NLGF}* module expression (from Fig. 3) is indicated in the top row and as a dashed line. Each diamond on the y-axis represents the donor's mean module expression across all cells (error bars: 95% confidence interval). Black diamonds indicate significant deviation from the *App^{Hu}* average (BH adjusted $p < 0.05$); grey diamonds: not significant. Density plots above each panel show the overall distribution of module z-scores across all *App^{Hu}* hMG.

Fig. S5. Plaque exposed hMG show enhanced MHCII expression in mouse xenograft and AD postmortem brain

(A) Immunofluorescence on the L23 line xenotransplanted in the *AppNLGF* mouse at 6 months of age, highlighting plaque pathology (X34, yellow), hP2RY12 (gray), hCD9 (cyan) and hMHCII (HLA-DR/DQ/DP, magenta). Scalebar = 50 μ m. (B) Chromogen immunohistochemical staining for MHCII (HLA-DR/DP/DQ, brown), and A β (4G8, red) in frontal cortex of symptomatic AD cases ($n=8$) and non-demented controls (non-AD, $n=7$). Representative images are shown for each group. Arrowheads: MHCII-positive microglia around A β plaques. Full arrows: MHCII-positive microglia not accumulating at an A β plaque.

(C) Quantifications are performed in 100 μ m circular regions of interest in areas containing A β plaques (AD, n=8), or no A β plaques (AD, n=8, and non-AD, n=7). The total number of MHCII-positive microglia per donor are summed up over all regions of interest. The Kruskal-Wallis test was performed followed by Dunn's multiple comparisons test (p-values indicated on graph). (D) Chromogen immunohistochemical staining for MHCII (HLA-DR/DP/DQ, brown), and CD4 T cells (red) in frontal cortex and white matter of a symptomatic AD case, showing little parenchymal CD4 T cells in AD brain. Full arrows: CD4-positive T cell in the extravascular space.

Fig. S6. Activation of allogenic naïve T cells by amyloid plaque exposed H9-hMG *ex vivo*

(A) Purity of naïve T cell populations, as measured by CD45RA-positivity, before and after enrichment for CD4⁺ and CD8⁺ T cells. (B) Experimental overview of the *ex vivo* hMG & human T cell co-culture experiment, where H9-derived xenotransplanted hMG are isolated from either the *App^{Hu}* or the *App^{NLGF-FIRE}* brain and co-cultured in a mixed leukocyte reaction with PBMC-isolated naïve T cells (1:5 ratio), after which the T cells are assessed for cytokine production and activation & proliferation. (C-D) Proliferation of CD45RO^{pos} CD8^{pos} T cells (C) and CD45RO^{pos} CD4^{pos} T cells (D) after 5-day co-culture with hMG from *App^{Hu}* (n=5) and *App^{NLGF-FIRE}* (n=6) mouse brain using Cell Trace Violet (CTV). Left panel: gating strategy; right panel: normalized quantification (to *App^{Hu}*). Unpaired parametric t-test, p-value indicated on graph. (E-F) Production of TNF α , IFN γ and IL-2 by CD8^{pos} T cells (E) and CD4^{pos} T cells (F) after 5-day co-culture with hMG from *App^{Hu}* (n=5) and *App^{NLGF-FIRE}* (n=6) mouse brain and stimulation with Cytostim. Left panel: gating strategy; right panel: normalized quantification (to *App^{Hu}*). Unpaired Mann-Whitney U test (E); unpaired t-test with Welch's correction (F); p-values indicated per graph.

Fig. S7. Donor-specific induction of WGCNA modules in response to A β pathology

Z-score shifts in module expression for each donor relative to the average *App^{NLGF}* amyloid response (dashed line, and "NLGF" diamond (red when significant) in top row, derived from Fig.3) and the donor's baseline module expression in *App^{Hu}* controls (solid black line at x=0). Each panel represents one WGCNA module. X-axis indicates z-score shift in module expression (1 unit = 1 standard deviation); Y-axis shows each donor's mean module score across all hMG (error bars = 95% confidence interval). Black diamonds: donors with significant deviation from the average *App^{NLGF}* response (BH adjusted $p < 0.05$); grey diamonds: not significant. Top density plot shows the overall distribution of module z-scores across all hMG from *App^{NLGF}*.

Fig. S8. Donor-specific transcriptional profiles are preserved in village versus single-donor grafts

(A) Correlation between differentially expressed gene profiles for each donor line, either xenotransplanted individually or as part of a hMG village. Differential expression was calculated relative to the donor L235 in both conditions. Correlation for L234 was low due to low cell numbers (<500) in the Village condition (see Methods).

Fig. S9. PRS correlations

(A) Correlation between PRS excluding the APOE locus (p-value threshold (pT)<0.05) and the Magenta module score in the donor-specific response to amyloid pathology with the donors from the HipSci cohort (blue, R=0.71, p=0.17) showing the same tendency as Fig.5C.

(B) Pearson correlation (R=0.26, p=0.17) between PRS (excluding APOE locus, p-value threshold (pT)<0.05) and the Magenta module score in the donor-specific profile in the physiological brain environment (hMG from *App^{Hu}* mice).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplTablesNatureSierksmaMoecharsBorrie.xlsx](#)
- [20251002PRSFigsNaturefinal.pdf](#)