Supplementary figures

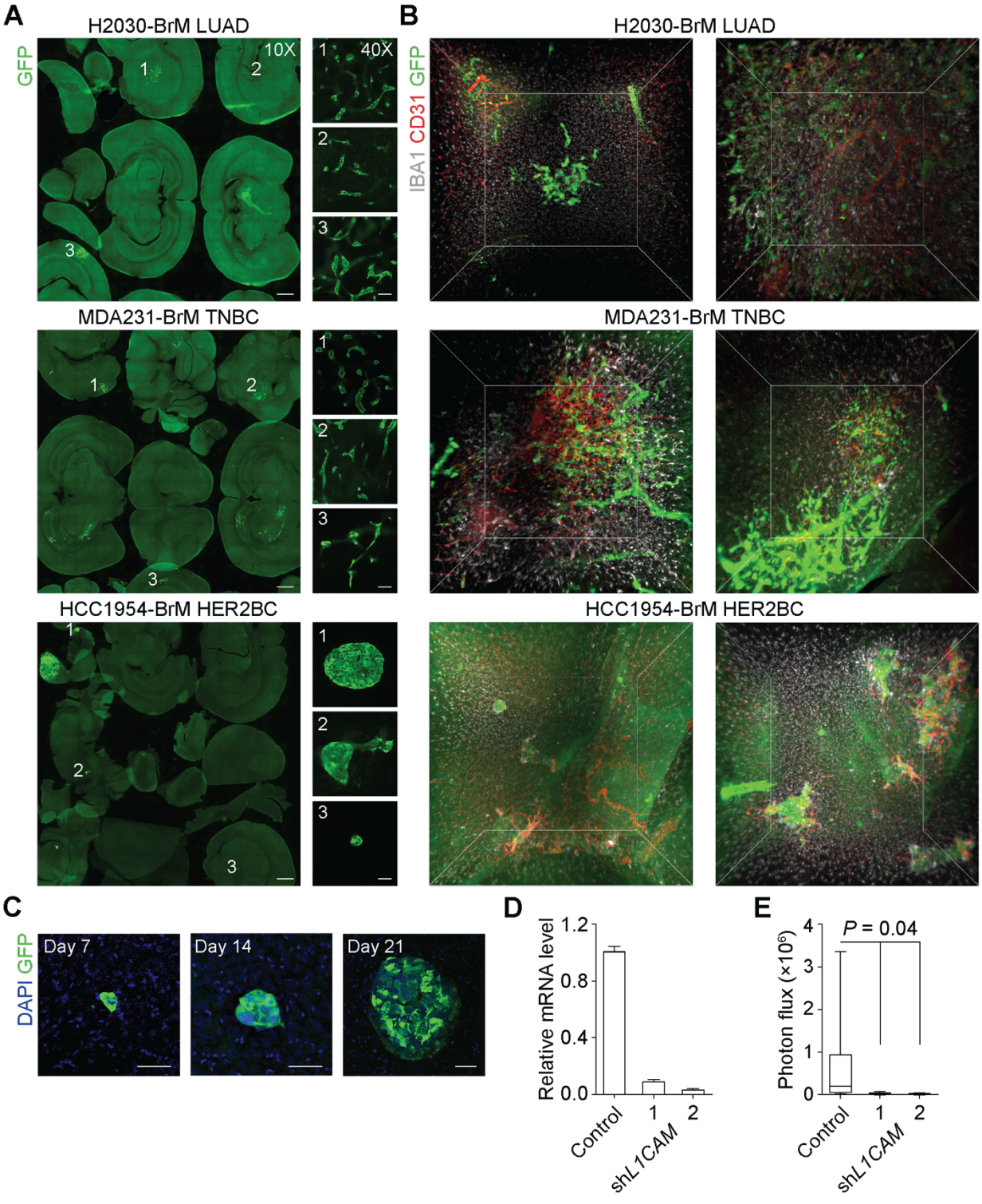


Figure S1. Brain colonization patterns and stromal interface examined by multiple imaging methods and impact of L1CAM on the survival of extravasated HER2BC cells. (A) Representative IF staining of free-floating tissue sections (80 μm) of athymic mouse brains harboring metastases formed by H2030-BrM LUAD, HCC1954-BrM HER2BC, and MDA231-BrM TNBC cells 3-4 weeks post-intracardiac inoculation. Scale bars, (left panels) 1 mm and (right panels) 50 μm in images obtained by 10X and 40X objectives, respectively. The same lesions imaged by both 10X and 40X objectives are indicated by matching numbers, with three example lesions shown for each model. (B) Representative 3D still images of immunolabeling-enabled imaging of solvent-cleared organs (iDISCO) showing cubic brain tissue regions (1 mm^3) harboring metastases 3-4 weeks post-intracardiac inoculation. Images were obtained by cleared-tissue lightsheet microscopy (Luxendo MuVi SPIM) imaging of whole-mount immunolabeling of a brain hemisphere. Corresponding animation of these still images is shown in Supplementary videos 1-3. (C) Representative IF staining of spheroidal HCC1954-BrM colonies formed at different times after intracardiac inoculation. Scale bars, 50 μm. (D) Relative mRNA levels of *L1CAM* in HCC1954-BrM cells expressing control vector or the indicated 2 shRNAs, measured by qRT-PCR. Mean ± SEM. (E) Effect of *L1CAM* shRNA knockdown (KD) in HCC1954-BrM cells on brain colonization, measured by *ex vivo* BLI of the brain (top panel, n = 6-7 mice/group, 4 weeks post-intracardiac inoculation).

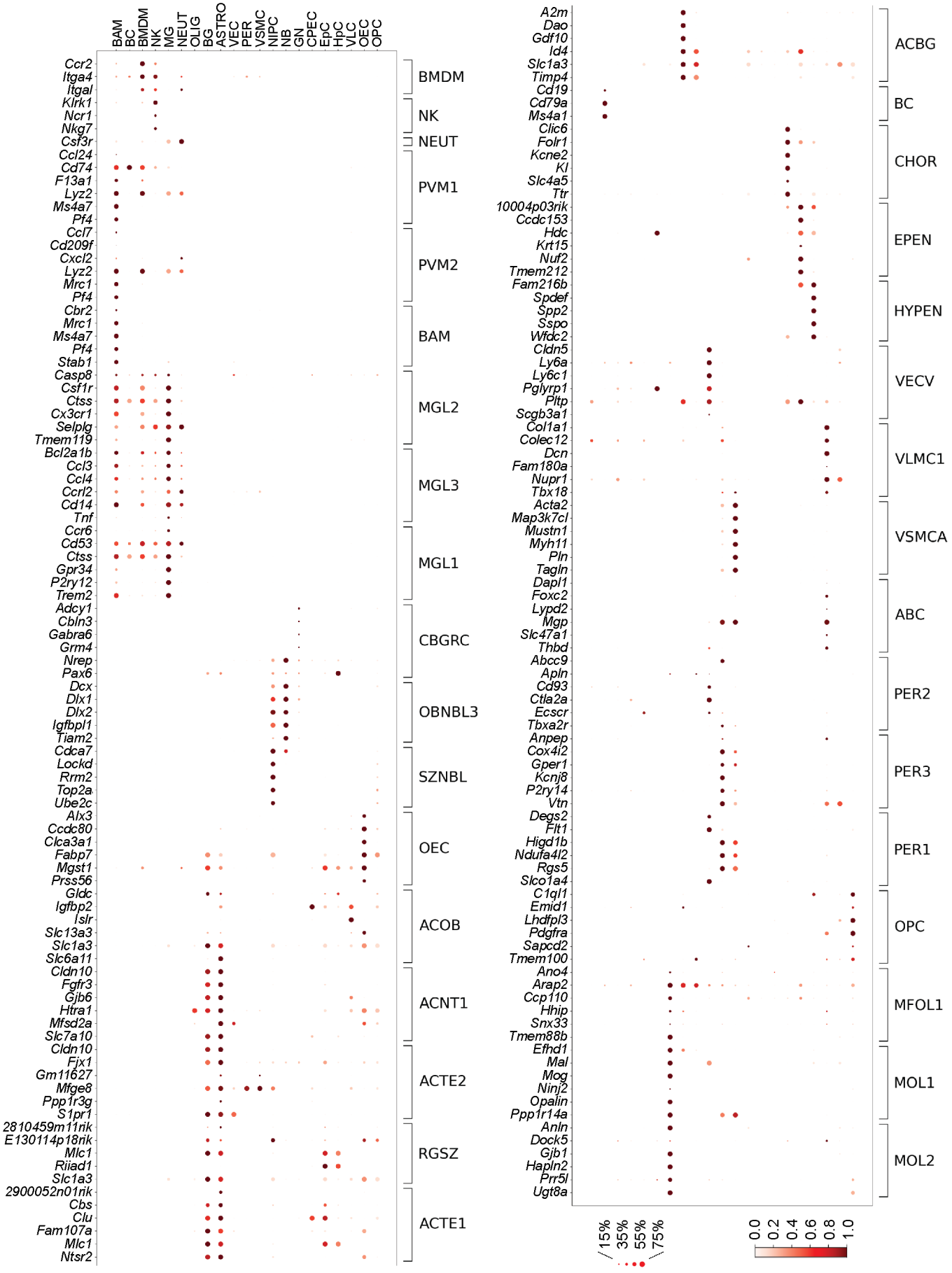


Figure S2. Expression of cell type marker genes in non-cancer cell populations from scRNA-seq profiling. Dot plots showing the expression of marker genes (rows) in each population of non-cancer cells (columns), computed on cells from both scRNA-seq experiment 1 and 2. The size of dots represents the fraction of cells with non-zero expression in each population; and saturation of color indicates the population-averaged log-transformed normalized UMI (unique molecular identifier) counts for each gene standardized to between 0 and 1. The cell populations are shown with identical abbreviated terms in Figure 2D, and in the same order with Figure 2E clustermap and 2F heatmaps (rows). Marker genes were obtained as described in Figure 2.

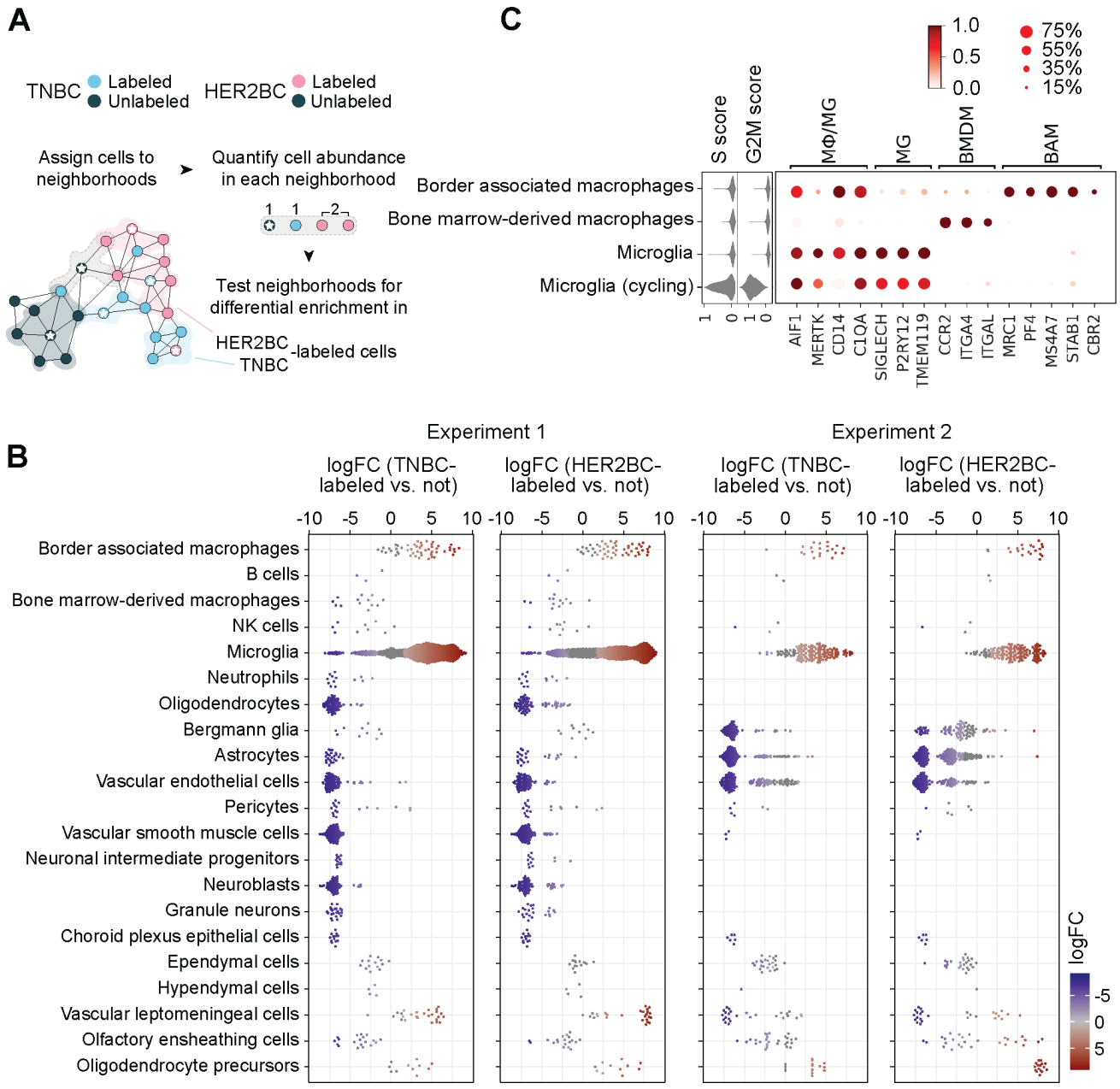


Figure S3. Microglia populations in the TME of MDA231 TNBC and HCC1954 HER2BC brain metastases. (A) Schematic of the statistical framework for quantifying and testing the differential abundance in cells from different sources (adapted from Ref.[47]). Cells (dots) are assigned to partially overlapping phenotypic neighborhoods (encircled by varied background colors, indicative of whether and which source of cells a neighborhood is enriched in) on a *k*-nearest neighbor (*k*-NN) graph. The neighborhoods are defined on index cells (indicated by white stars), selected using a graph sampling algorithm. The counts of cells in the light gray

neighborhood (not enriched in cells from any particular source) are shown as an example. (B) Differential abundance (measured by log fold change, logFC, x axis) of TNBC-labeled (left panels) and HER2BC-labeled (right panels) cells in reference to unlabeled cells among all phenotypic neighborhoods, stratified by the types of neighborhood index cells. Results of experiment 1 and 2 are presented in parallel. Phenotypic neighborhoods with significant differential abundance (BH-adjusted P values ≤ 0.1) are colored to indicate corresponding logFC (x axis) values. (C) Expression of canonical cell type markers (shown by dot plot, right panel) and S phase and G2/M phase scores (violin plot, left panel) in the indicated subsets of macrophages. MΦ, macrophages. MG, microglia computationally identified to be in the G1 phase or not cycling. MG (cycling), microglia inferred to be cycling, given high scores of the S and G2/M phases. BAM, border associated macrophages. BMDM, bone marrow-derived macrophages. The size of dots in the dot plot represents the fraction of cells with non-zero expression in each subset of cells; and saturation of color indicates the subset-averaged log-transformed normalized UMI (unique molecular identifier) counts for each gene standardized to between 0 and 1.

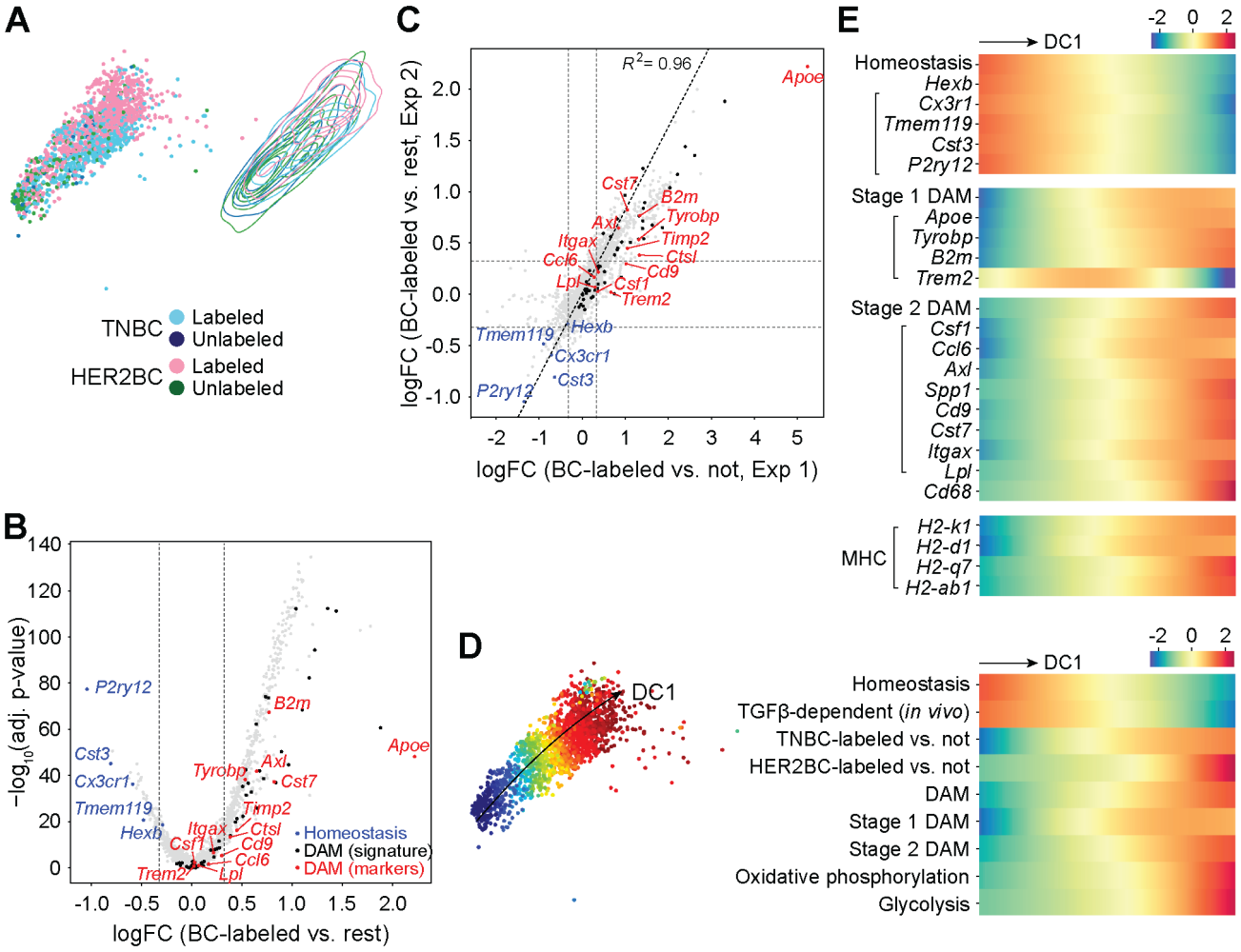


Figure S4. Replication of homeostasis-to-DAM transition in metastasis-associated microglia by an independent scRNA-seq profiling experiment. All results computed on the non-cycling microglia (MG) from experiment 2. (A) (Left panel) UMAP embedding of the 4 indicated sources of altogether 1737 cells, and (right panel) contour plots of each source in the embedding, as in Figure 4A. (B) Volcano plot of the logFC in gene expression against corresponding BH-adjusted P values, comparing the groups of phenotypic neighborhoods concordantly enriched in breast cancer (BC) cell-labeled non-cycling microglia to the rest. 501 (92%) of 546 DEGs detected overlap with the 868 DEGs from experiment 1. (C) Linear regression showing that the logFC values of all 11379 genes with estimable expression coefficients in metastasis-associated non-cycling microglia were well correlated between experiments 1 and 2 (see Table S3 and STAR Methods). x- and y-axis values correspond to the x-axis values in Figure 3C and Figure S4B, respectively. (B, C) As with Figure 3C, logFC thresholds were indicated by gray dashed lines, and

sources and lists of indicated gene sets were provided in Table S2. (D) (Left panel) UMAP plot showing first diffusion component (DC1) values computed on all cells by color map; (right panel) heatmap showing fitted trends of indicated neighborhood-level signature scores and labeled cell enrichment (quantified by the local logFC in abundance in reference to unlabeled cells, logFC) along the DC1 values of neighborhood index cells (rows, z-normalized per row across neighborhoods). Marker genes were obtained as detailed in Figure 4F. (E) Heatmap showing fitted trends of neighborhood-level signature scores and gene expression along the DC1 values of neighborhood index cells as shown in the heatmap in (B) (rows, z-normalized per row). Trends of inflammatory cytokine genes *Tnf*, *Ccl4*, *Il1b* expression were not computed given their overall low UMI counts, possibly caused by the addition of transcription and translation inhibitors during the brain tissue harvesting and homogenizing steps in experiment 2.

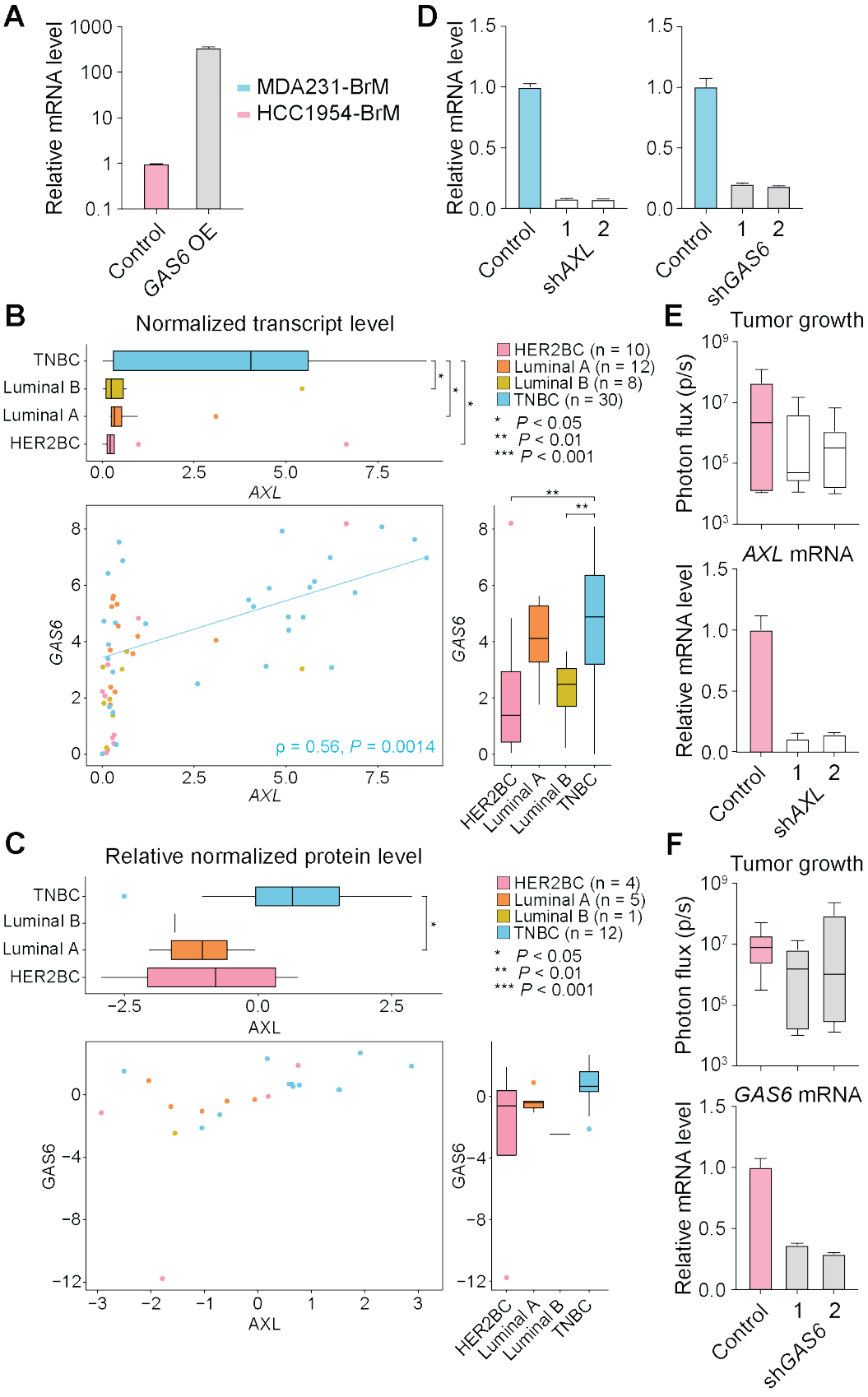
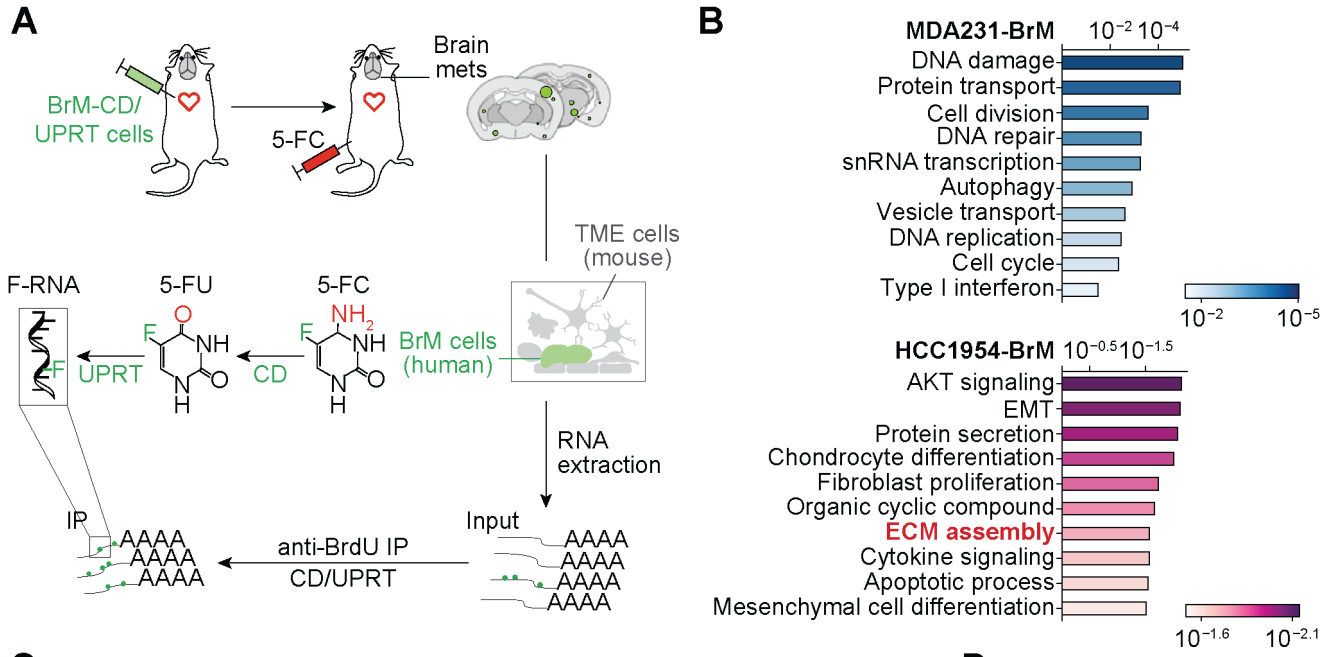
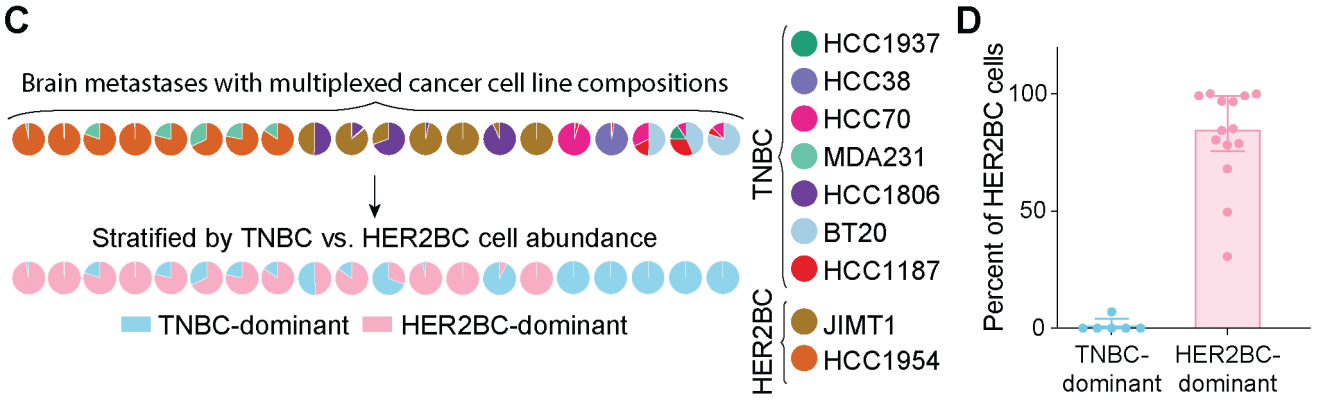


Figure S5. Perturbing GAS6/AXL signaling in MDA231 TNBC and HCC1954 HER2BC brain metastases. (A) Relative mRNA levels of *GAS6* in HCC1954-BrM cells overexpressing (OE) *GAS6* or control vector, measured by qRT-PCR. Mean ± SEM. (B-C) Co-enrichment of GAS6 and AXL in patient TNBC cells. Normalized transcript (B) and relative normalized protein (C) levels of AXL and GAS6 in all available human breast cancer cell lines in DepMap database (depmap.org/portal/). ρ, Spearman's correlation coefficient. (D) *AXL* (left panel) and *GAS6* (right panel) relative mRNA levels in MDA231-BrM cells expressing control vector or the indicated shRNAs (2 shRNAs per target gene), validated by qRT-PCR. Mean ± SEM. (E-F) Effect of *AXL* (E) and *GAS6* (F) shRNA knockdown (KD) in HCC1954-BrM cells (2 shRNAs per target gene), validated by qRT-PCR (bottom panels, mean ± SEM), on brain colonization measured by *ex vivo* BLI of the brain (top panels, n = 6-7 mice/group, 4 weeks post-intracardiac inoculation).





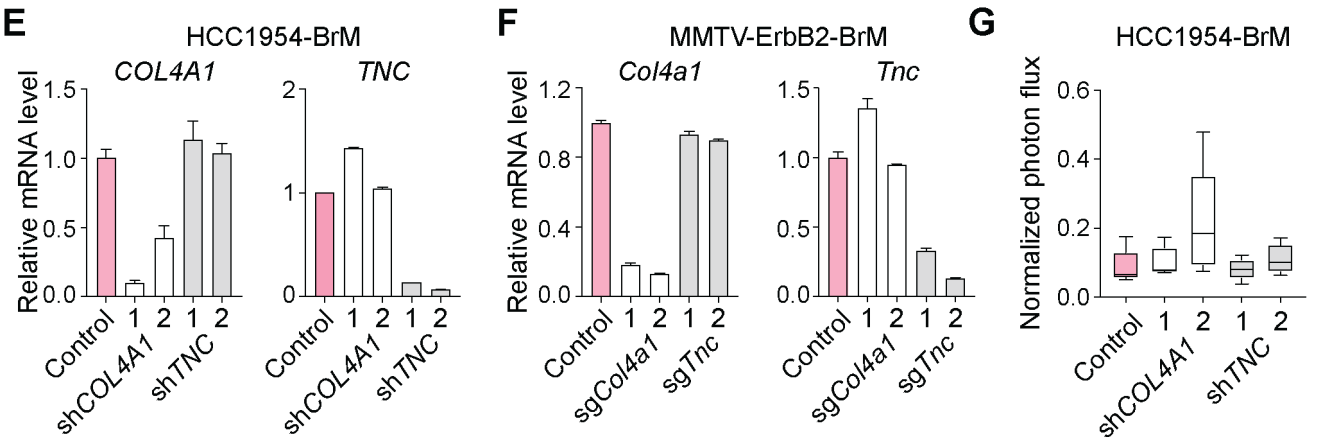


Figure S6. Characterizing and suppressing the expression of matrisome genes in TNBC and HER2BC brain metastases. (A) Schematic illustrating the Flura-seq profiling of *in situ* transcriptome of brain metastases. Athymic mice were inoculated intracardiacally with brain metastatic (BrM) cells engineered to co-express cytosine deaminase (CD) and uracil phosphoribosyl transferase (UPRT). Mice were administered with 5-fluorocytosine (5-FC) (250 mg/kg) by intraperitoneal (IP) injection 12 hours prior to harvesting the brain tissue. 5-FC was converted through a series of steps to 5-fluorouridine triphosphate (F-UTP) specifically in CD/UPRT expressing BrM cells. F-UTP was incorporated into nascent RNA of the cells to yield the 5-FU-tagged RNA (F-RNA) that could be purified by anti-BrdU immunoprecipitation for RNA-seq analysis. Specifically, CD converted 5-FC to 5-fluorouracil (5-FU); and UPRT converted 5-FU to cell membrane-impermeable 5-fluorouridine monophosphate (F-UMP), which was ultimately converted to 5-fluorouridine triphosphate (F-UTP) by endogenous enzymes. (B) Overrepresentation analysis showing the top gene ontology (GO) terms associated with genes respectively upregulated in MDA231 and HCC1954 brain metastases, compared in reference to each other (see STAR Methods). Color shades indicate BH-adjusted P values of normalized enrichment scores. (C) Composition of cells from indicated TNBC and HER2BC cell lines in brain metastases formed by each multiplexed cell line pool, quantified by deep sequencing of the barcodes of cell lines. (D) Percentage of HER2BC cells in multiplexed brain metastasis samples classified to be dominated by TNBC (< 25% HER2BC cells) and HER2BC cells, respectively. Median with interquartile range. (E) *TNC* and *COL4A1* relative mRNA levels in HCC1954-BrM cells expressing control vector or the indicated shRNAs (2 shRNAs per target gene), measured by qRT-PCR. Mean ± SEM. (F) *Tnc* and *Col4a1* relative mRNA levels in MMTV-ErbB2-BrM cells expressing control vector or the indicated CRISPRi sgRNA, measured by qRT-PCR. Mean ± SEM. **(**G) Effect of *TNC* and *COL4A1* shRNA knockdown (KD) on lung colonization by HCC1954-BrM cells. Lung photon flux was measured 4 weeks post-tail vein inoculation (n = 5 mice/group).

Supplementary tables and videos

Table S1. Reference marker genes of cell type and state clusters.

Table S2. Curated gene sets of microglia phenotypes.

Table S3. DEGs of breast cancer-labeled microglia.

Table S4. Microglia GSEA results with BP gene sets.

Table S5. Microglia GSEA results without BP gene sets.

Table S6. Upregulated DEGs of HER2BC-BrM.

Supplementary video 1. 3D animation of immunolabeling-enabled imaging of solvent-cleared organs (iDISCO) showing H2030-BrM metastases, corresponding to Figure S1B (first row).

Supplementary video 2. 3D animation of immunolabeling-enabled imaging of solvent-cleared organs (iDISCO) showing MDA231-BrM metastases, corresponding to Figure S1B (second row).

Supplementary video 3. 3D animation of immunolabeling-enabled imaging of solvent-cleared organs (iDISCO) showing HCC1954-BrM metastases, corresponding to Figure S1B (third row).