Supplementary Information

for Single-cell and spatial transcriptomics reveal a cyst–neuron–microglia network underlying brain damage during *Toxoplasma gondii* infection

Chi Zhang, Yingying Lu, Weiling Wu, Zhihao Zhou, Jiating Chen, Wenjun Huang, Min Chen, Weihao Zou, Lijuan Zhou, Shuyu Zheng, Xiangyi He, Xinyue Liao and Hongjuan Peng

Contents: Supplementary Note 1; Supplementary Methods 1; Supplementary Figs. 1–4; Supplementary Tables 1–7.

Supplementary Note 1 | Annotation of host cell subclusters

1.1 Neuronal subclusters (Supplementary Fig. 1b)

The 27 neuronal subclusters comprise cerebellar granule cells (C-GC), highly plastic spiny neurons (HPSN), olfactory bulb granule cells (OB-GC), tectal sensory glutamatergic projection neurons (TS-GPN), perirhinal glutamatergic projection neurons (PRh-GPN), association cortical pyramidal neurons of layers II–IV (AC-GPN-2/4), direct-pathway D1 receptor-positive striatal medium spiny neurons (D1-MSN), highly active spiny neurons (HASN), association cortical pyramidal neurons of layers II–III (AC-GPN-2/3), indirect-pathway D2 receptor-positive striatal medium spiny neurons (D2-MSN), dentate gyrus granule cells (DG-GC), association cortical pyramidal neurons of layers V–VI (AC-GPN-5/6), hippocampal CA1/CA3 pyramidal neurons (Hipp-GPN-CA1/3), SST^+ interneurons (SST^+ IN), basket cells (BC), neuroblasts (NB), periglomerular cells (PGC), VIP^+ GABAergic interneurons (VIP^+ IN), hybrid-phenotype young dopaminergic neurons (HP-yDA), short-axon cells (SAC), chandelier cells (ChC), nucleus accumbens-derived direct-pathway D1 medium spiny neurons with Drd3 co-expression ($Drd3^+$ NAc_D1-MSN), Purkinje cells (PkC), mitral cells (MC), circuit-integrating neurons (CIN), thalamic glutamatergic projection neurons (Thal-GPN) and glycinergic interneurons (GlyIN). Glutamatergic neurons such as cortical pyramidal neurons (AC-GPNs-2/3, AC-GPNs-2/4, AC-GPNs-5/6) and hippocampal pyramidal neurons (Hipp-GPNs-CA1/3) decreased significantly in proportion, as did GABAergic neurons such as striatal D1-MSNs, D2-MSNs and VIP^+ INs, and olfaction-related cells (OB-GCs, PRh-GPNs, PGCs, SACs, MCs). The marked decline of neuroblasts (NBs) and immature neurons (HP-yDA) suggests that *T. gondii* infection may impair adult neurogenesis. The two functionally defined neuronal subclusters identified by unsupervised clustering, highly active spiny neurons (HASN) and highly plastic spiny neurons (HPSN), showed significant increase and decrease, respectively, indicating reduced neuronal plasticity upon infection; the increased proportion of HASNs during chronic infection may relate to parasite-promoted neuronal dopamine production. The increase in circuit-integrating neurons (CINs) may suggest that infection promotes neuronal circuit re-integration.

1.2 Astrocyte subclusters (Supplementary Fig. 1c)

The eight astrocyte subclusters comprise homeostatic astrocytes 1–3 (HA1, HA2, HA3), interferon-responsive astrocytes (IRA), transitional reactive astrocytes (TRA), immune effector astrocytes (IEA), vascular-associated astrocytes (VA) and oxidative stress-responsive astrocytes (ORA).

1.3 Microglial subclusters (Supplementary Fig. 1d)

The five microglial subclusters comprise interferon-responsive microglia (IRM), homeostatic microglia 1 and 2 (HM1, HM2), proliferative microglia (PM) and pro-inflammatory microglia (PIM). As brain-resident immune cells, microglia respond rapidly to brain injury or pathogen infection. Microglial numbers increased significantly during acute *T. gondii* infection and remained elevated through the chronic phase. Homeostatic microglia (HM1 and HM2) decreased significantly after infection, whereas IRM, PM and PIM increased; homeostatic microglial numbers partially recovered during chronic infection. Pseudotime analysis showed that early-pseudotime cells were mainly homeostatic microglia, predominantly from uninfected mice; as infection proceeded, homeostatic microglia differentiated into IRM, PM and PIM pro-inflammatory/proliferative states.

1.4 Monocyte/macrophage subclusters (Supplementary Fig. 1e)

The five monocyte/macrophage subclusters comprise classical monocytes (cMo), $Tgfbr1^+$ monocyte-derived macrophages ($Tgfbr1^+$ MoDM), M1 monocyte-derived macrophages (Mo-M1 Mac), regulatory monocyte-derived macrophages (Mo-Mreg) and border-associated macrophages (BAM). Among peripheral immune cells, monocytes/macrophages are rapidly recruited to the CNS and infiltrate the brain parenchyma after *T. gondii* infection; brain-resident border-associated macrophages also contribute to pathogen defense. Driven by peripheral and BBB-associated pro-inflammatory cues and

by direct parasite stimuli, circulating monocytes progressively differentiate into macrophages that control *T. gondii* infection through antigen presentation and pro-/anti-inflammatory functions.

1.5 Dendritic cell subclusters (Supplementary Fig. 1f)

The four dendritic cell subclusters comprise conventional dendritic cells type 1 (cDC1) and type 2 (cDC2), migratory dendritic cells (migDC) and plasmacytoid dendritic cells (pDC).

1.6 T cell subclusters (Supplementary Fig. 1g)

The nine T cell subclusters comprise $CD8^+$ T cells 1 ($CD8^+$ T 1), $CD4^+$ T cells/T follicular helper cells ($CD4^+$ T/Tfh), activated $CD8^+$ cytotoxic T cells/natural killer T cells (Act $CD8^+$ T/NKT), exhausted/proliferating $CD8^+$ T cells (Tex/Tpro), γδT cells (γδT), $CD8^+$ T cells 2 ($CD8^+$ T 2), regulatory T cells (Treg), T helper 2 cells (TH2) and $IL7R^+$ memory T cells ($IL7R^+$ Tm). Among peripheral immune cells, T cells are rapidly recruited to the CNS and infiltrate the brain parenchyma after *T. gondii* infection. T cells are extremely rare in uninfected mouse brain, increase markedly during acute and chronic infection, and differentiate into functionally distinct subclusters, further supporting their important role in anti-*T. gondii* immunity during acute and chronic infection.

1.7 B cell and plasma cell subclusters (Supplementary Fig. 1h, i)

The four B cell subclusters comprise naive B cells (Naive B), activated B cells (Act B), germinal center B cells (GC B) and memory B cells (Mem B). The two plasma cell subclusters comprise mature plasma cells (MPC) and plasmablasts (PB). Unlike T cells, B cells and plasma cells were not recruited to the brain during acute infection: they were identified only in chronically infected brains, and plasma cells were entirely absent from uninfected mice. Beyond producing antibodies against *T. gondii*, B cells have recently been shown to maintain the architecture of the antigen-presenting cell network, indirectly determining the efficiency of $CD8^+$ T cell priming, which may partly explain their marked increase during chronic *T. gondii* infection.

1.8 Fibroblast subclusters (Supplementary Fig. 1j)

The five fibroblast subclusters comprise myofibroblasts (MFB), lymphocyte-interacting fibroblasts (LIFB), dural fibroblasts (DFB), inflammatory fibroblasts (IFB) and leptomeningeal fibroblasts (LFB).

1.9 Mural/endothelial cell subclusters (Supplementary Fig. 1k)

The three mural/endothelial subclusters comprise endothelial cells (EC), pericytes (Peri) and vascular smooth muscle cells (VSMC).

Supplementary Methods 1 | LC–MS mobile-phase gradient and mass-spectrometry parameters

Bulk metabolomics was performed on a Waters ACQUITY Premier ultra-high-performance liquid chromatography (UPLC) system coupled to a quadrupole time-of-flight mass spectrometer, equipped with an ACQUITY Premier BEH C18 1.7 µm VanGuard FIT column (2.1 × 100 mm). Phase A was 60% acetonitrile/40% water and phase B was 90% isopropanol/10% acetonitrile, both containing 10 mM ammonium formate and 0.1% formic acid. The mobile-phase gradient is given in Table 1 and the mass-spectrometry parameters in Table 2.

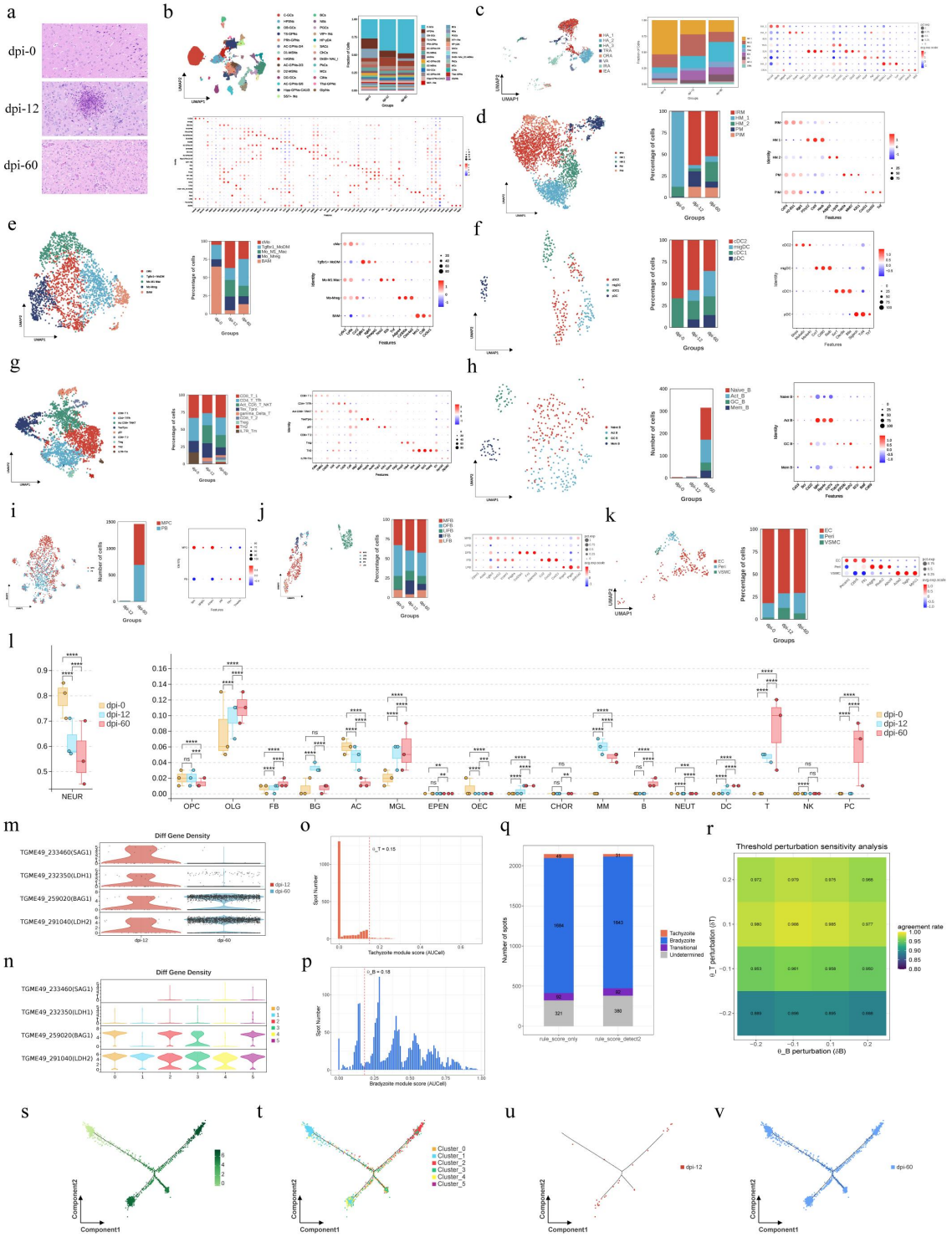
Table 1 | Mobile-phase gradient

Time (min)	Flow rate (mL/min)	Phase A (%)	Phase B (%)	Curve
Initial	0.2	60	40	6
4.00	0.2	57	43	6
4.10	0.2	50	50	6
24.00	0.2	46	54	6
24.10	0.2	30	70	6
36.00	0.2	1	99	6
40.00	0.2	1	99	6
40.10	0.2	60	40	6
44.00	0.2	60	40	6

Table 2 | Mass-spectrometry parameters

Parameter	Value
Capillary voltage	2.30 kV
Cone voltage	40 V
Source offset	10 V
Source temperature	150 °C
Desolvation temperature	500 °C
Cone gas flow	50 L/h
Desolvation gas	6.0 bar
Reference capillary voltage	2.20 kV
Acquisition mode	MSE
Low collision energy	6 eV
High collision energy	30–50 eV
Lock mass	Leucine enkephalin, 0.2 µg/mL
Mass-axis calibration	Sodium formate, 0.5 mmol/L

Supplementary Fig. 1



Supplementary Fig. 1 | Host cell subclustering, parasite marker-gene expression and stage-classification quality controls.

a, H&E staining of brain sections at 0, 12 and 60 dpi.

b–k, UMAP, stage composition and marker-gene dot plots for major host populations: neurons (b), astrocytes (c), microglia (d), monocytes/macrophages (e), dendritic cells (f), T cells (g), B cells (h), plasma cells (i), fibroblasts (j), and mural/endothelial cells (k). Annotations are in Supplementary Note 1.

l, Major cell-population frequencies across stages (n = 3 mice per time point; one-way ANOVA with Tukey's test; ns, not significant; *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001).

m,n, SAG1, LDH1, BAG1 and LDH2 expression by infection stage (m) and parasite subcluster (n).

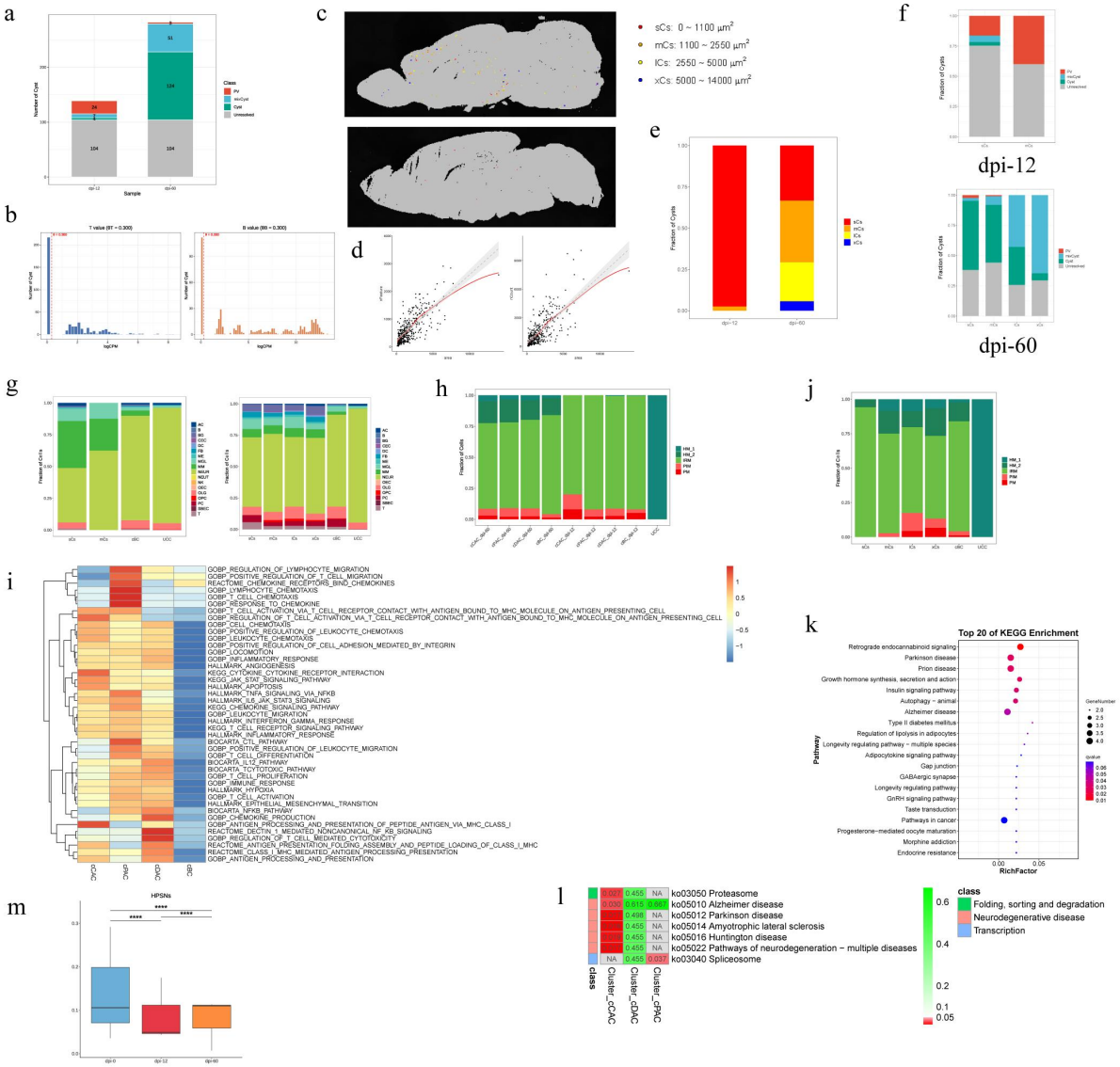
o,p, Tachyzoite (o) and bradyzoite (p) AUCell-score distributions for 2,146 parasites. Dashed lines mark $\theta_T = 0.15$ and $\theta_B = 0.18$; threshold determination is described in Methods.

q, Stage-class counts using score thresholds alone or with an additional requirement of ≥ 2 detected genes per active module (score-only versus detection rule: Bradyzoite, 1,684 versus 1,643; Tachyzoite, 49 versus 31; Transitional, 92 versus 92; Undetermined, 321 versus 380; Supplementary Table 1).

r, Agreement of stage assignments across 16 combinations of ±10% and ±20% threshold perturbations (agreement, 88.8–98.8%; Supplementary Table 6).

s–v, Monocle 2 pseudotime trajectories of six parasite subclusters, coloured by pseudotime (s), subcluster (t), and cells from 12 dpi (u) and 60 dpi (v).

Supplementary Fig. 2



Supplementary Fig. 2 | Stage and size classification of PV/cyst structures and cyst-neighbourhood composition.

a, Stage classification of parasite foci. Of 440 delineated foci, 19 without parasite UMIs were excluded, leaving 421 (139 at 12 dpi; 282 at 60 dpi). Tachyzoite (T) and bradyzoite (B) logCPM module scores ($\theta_T = \theta_B = 0.30$) defined PV ($T^+B^-$), Cyst ($T^-B^+$), mixCyst ($T^+B^+$) and Unresolved. At 12 dpi, counts were PV 24, mixCyst 7, Cyst 4 and Unresolved 104; at 60 dpi, PV 3, mixCyst 51, Cyst 124 and Unresolved 104. PVs comprised 24 of 35 classifiable 12-dpi foci, and cysts (Cyst plus mixCyst) 175 of 178 classifiable 60-dpi foci (Supplementary Table 2).

b, T and B module-score distributions across 421 foci; dashed lines mark $\theta_T = \theta_B = 0.30$.

c, Spatial distributions of size classes (sCs, mCs, lCs, xCs) at 12 and 60 dpi.

d, Detected transcripts (nCount) and genes (nFeature) per structure versus cross-sectional area, with linear regression fits (nCount–area: R = 0.722, P = 5.84×10^{-62}; nFeature–area: R = 0.745, P = 6.47×10^{-68}).

e,f, Size-class composition by infection stage (e) and stage composition within each size class (f). At 60 dpi, mixCyst accounted for 4%, 13%, 58% and 92% of classifiable sCs, mCs, lCs and xCs, respectively.

g, Host cell-type composition of CN by size class, with cBC and UCC as references.

h, Microglial subcluster composition across CN layers at 12 and 60 dpi.

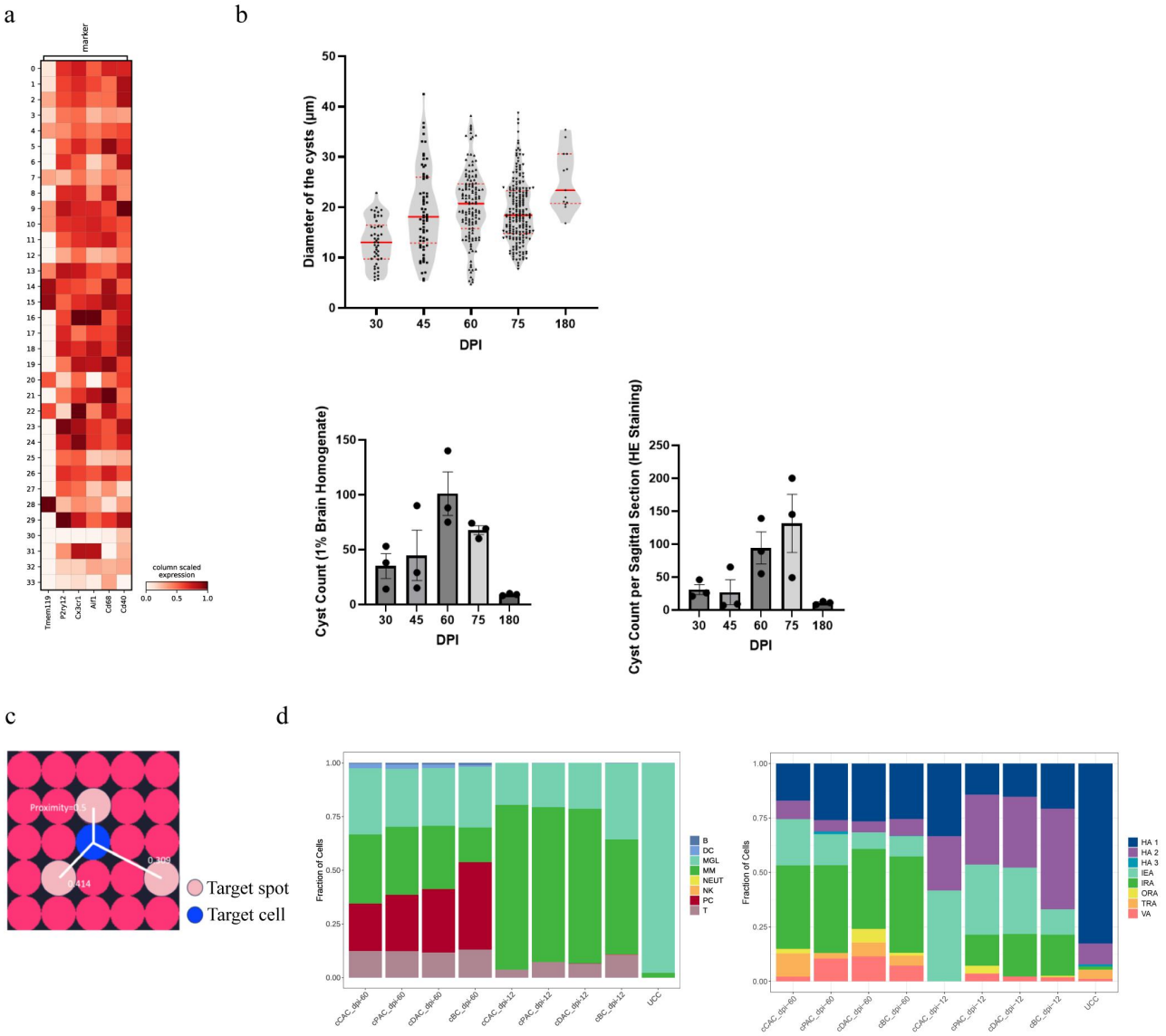
i, GSVA of cytotoxic $CD8^+$ T-cell chemotaxis, activation and differentiation pathways in CN-associated IRMs across layers and cBC.

j, Microglial subcluster composition by structure size class.

k,l, KEGG enrichment for DEGs from microglia associated with large structures (k, top 20 terms) and across CN layers (l); dot size, gene number; colour, q value.

m, HPSN frequency across infection stages (n = 3 mice per time point; one-way ANOVA with Tukey's test; $P < 0.001$, $P < 0.0001$).

Supplementary Fig. 3



Supplementary Fig. 3 | Stereo-cell recovery controls, cyst-burden kinetics, proximity-score schematic and CN composition.

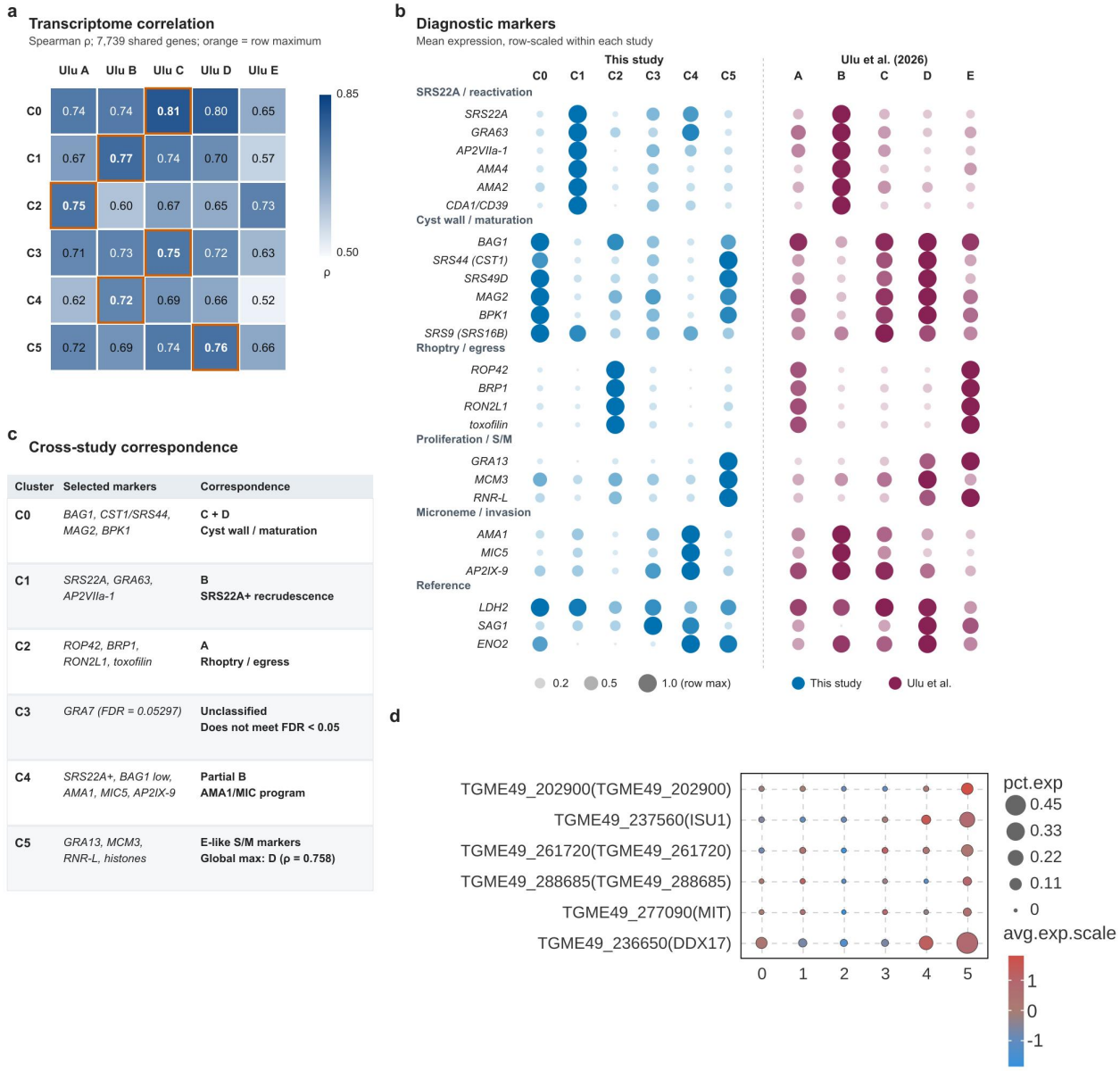
a, Marker-expression heatmap (column-scaled) of cells recovered from Stereo-cell chips after the 6-h co-culture (microglial markers Tmem119, P2ry12, Cx3cr1, Aif1, Cd68, Cd40); rows are the recovered on-chip cell clusters (0–33), confirming that nearly all recovered cells are primary microglia.

b, Cyst-burden kinetics in the 129S2/ME49 oral-infection model: cyst diameter (top; violins with individual cysts as dots; red solid lines, medians; red dashed lines, quartiles) and cyst counts (bottom left, per 1% brain homogenate; bottom right, per sagittal H&E section; bars show mean ± s.e.m., individual mice shown as dots) at 30, 45, 60, 75 and 180 dpi (n = 3 mice per time point).

c, Schematic of the proximity score used to define BBB adhesion: the score between a query spot and target spots is computed from their spatial relationship (illustrated values 0.5, 0.414 and 0.309); peripheral immune cells with proximity to endothelial cells > 0.414 were defined as BBB-adherent.

d, Composition of immune cell populations (left) and astrocyte activation states (HA1, HA2, HA3, IRA, TRA, IEA, VA, ORA; right) across CN layers at 12 and 60 dpi, with cBC and UCC as references, showing activation of CN-enriched immune and atypical immune cells.

Supplementary Fig. 4



Supplementary Fig. 4 | Cross-study comparison of bradyzoite subtypes and stress/adaptation marker expression.

a, Spearman correlation of mean expression over 7,739 shared genes between six parasite subclusters (C0–C5) and five in vivo bradyzoite groups (A–E) from Ulu et al. Orange frames denote the highest correlation in each row.

b, Mean expression of 25 diagnostic markers, scaled to the maximum for each gene within each study. Circle size and opacity denote the fraction of the row maximum. Modules group markers by biological program.

c, Cross-study correspondence based on global correlations and diagnostic markers. C5 shows E-like S/M-marker expression despite its highest global correlation with Ulu D ($\rho = 0.758$); C4 partially overlaps with Ulu B but adds a microneme/AMA1 program; C3 is unclassified because its only candidate marker, GRA7, does not meet FDR < 0.05 (FDR = 0.05297).

d, Dot plot of crowding-stress- and antioxidant-related transcripts (TgPRO (TGME49_202900), TgISU1 (TGME49_237560), TGME49_261720 (ZIP-family metal cation transporter), HscB (TGME49_288685; Fe–S protein assembly co-chaperone), TgMIT (TGME49_277090) and TgDDX17 (TGME49_236650)) across parasite subclusters 0–5. Dot size indicates the percentage of expressing cells (pct.exp); colour indicates scaled mean expression (avg.exp.scale).

Supplementary Tables

Full data are provided in the accompanying Excel files. Each workbook includes a Table legend worksheet containing its title and description.

Supplementary Table 1 | Per-cell tachyzoite/bradyzoite module scores and stage classification of intracerebral parasites.

Sheet 1 (Per-cell module scores): for each of the 2,146 analysed parasite cells: cell ID, tachyzoite- and bradyzoite-module AUCell scores, sample, subcluster (0–5), time point (12 or 60 dpi), and the stage calls under the score-only rule (rule_score_only) and the rule additionally requiring ≥ 2 detected module genes (rule_score_detect2). Sheet 2 (Classification rule counts): class counts under each rule. Data correspond to Fig. 1g, h and Supplementary Fig. 1m–r.

Excel file: Suppl_Table1_per_cell_classification.xlsx

Supplementary Table 2 | Per-focus module scores and stage classification of spatial PV/cyst structures.

Sheet 1 (Per-focus classification): for each of the 421 analysed parasite foci: focus ID, tachyzoite (T) and bradyzoite (B) module scores (mean within-parasite logCPM), time point, cross-sectional area, numbers of detected tachyzoite (nzT) and bradyzoite (nzB) panel genes, and the assigned class (PV, mixCyst, Cyst, Unresolved). Sheet 2 (Focus class counts): class counts per time point. Data correspond to Supplementary Fig. 2a.

Excel file: Suppl_Table2_per_focus_classification.xlsx

Supplementary Table 3 | Differentially expressed genes between cyst-neighbourhood microglia (CNM) and non-neighbourhood microglia (NNM).

The worksheet (CNM vs NNM DEGs) lists genes detected as differentially expressed between CNM and NNM recovered from Stereo-cell chips (Wilcoxon rank-sum test; |log2 fold change| > 0.5 and P < 0.05). Columns: gene symbol, Wilcoxon score, log fold change, raw P value, Benjamini–Hochberg-adjusted P value, absolute log2 fold change and Ensembl gene ID. Data correspond to Fig. 3h, i.

Excel file: Suppl_Table3_CNM_vs_NNM_DEG.xlsx

Supplementary Table 4 | Differential metabolites from bulk and spatial metabolomics.

Metabolites with VIP > 1. Worksheets: NEG UP, bulk negative-ion upregulated; NEG DOWN, bulk negative-ion downregulated; POS UP, bulk positive-ion upregulated; POS DOWN, bulk positive-ion downregulated; SM UP, spatial cyst-neighbourhood upregulated; SM DOWN, spatial cyst-neighbourhood downregulated. Columns are Compound ID (HMDB/LIPID MAPS where available) and Description. Data correspond to Fig. 5c–h; per-feature VIP scores and coefficients are provided in Supplementary Table 7.

Excel file: Suppl_Table4_differential_metabolites.xlsx

Supplementary Table 5 | Tachyzoite and bradyzoite stage-module genes.

The single worksheet (Stage module genes) lists the 14 genes used for tachyzoite/bradyzoite stage classification (7 genes per module). Columns: module assignment, gene symbol, synonyms, ToxoDB gene ID (ME49) and protein description. These gene panels were used for AUCell-based classification of parasite single cells and for spatial PV/cyst focus classification as described in Methods.

Excel file: Suppl_Table5_stage_module_genes.xlsx

Supplementary Table 6 | Sensitivity of parasite stage classification to threshold perturbations.

The worksheet (Threshold sensitivity) reports 16 combinations of relative perturbations (dT and dB: −0.2, −0.1, +0.1 and +0.2) applied to the adopted tachyzoite and bradyzoite thresholds (thetaT and thetaB), and the resulting agreement rate (agree_rate) with the reference classification. The adopted thresholds were thetaT = 0.15 and thetaB = 0.18; agreement ranged from 88.8% to 98.8%. Data correspond to Supplementary Fig. 1r.

Excel file: Suppl_Table6_threshold_sensitivity.xlsx

Supplementary Table 7 | OPLS-DA coefficients and VIP scores for metabolomics features.

Worksheets: Bulk NEG, bulk brain negative-ion mode (112 features); Bulk POS, bulk brain positive-ion mode (45 features); Spatial SM, spatial cyst-neighbourhood negative-ion mode (183 features). Columns: m/z, OPLS-DA coefficient (CoeffCS), VIP, Compound ID (HMDB/LIPID MAPS, where available) and Description. The VIP = 1 threshold was used to define differential metabolites (Supplementary Table 4). Data correspond to Fig. 5c–h.

Excel file: Suppl_Table7_VIP_analysis.xlsx