Supplementary Figures

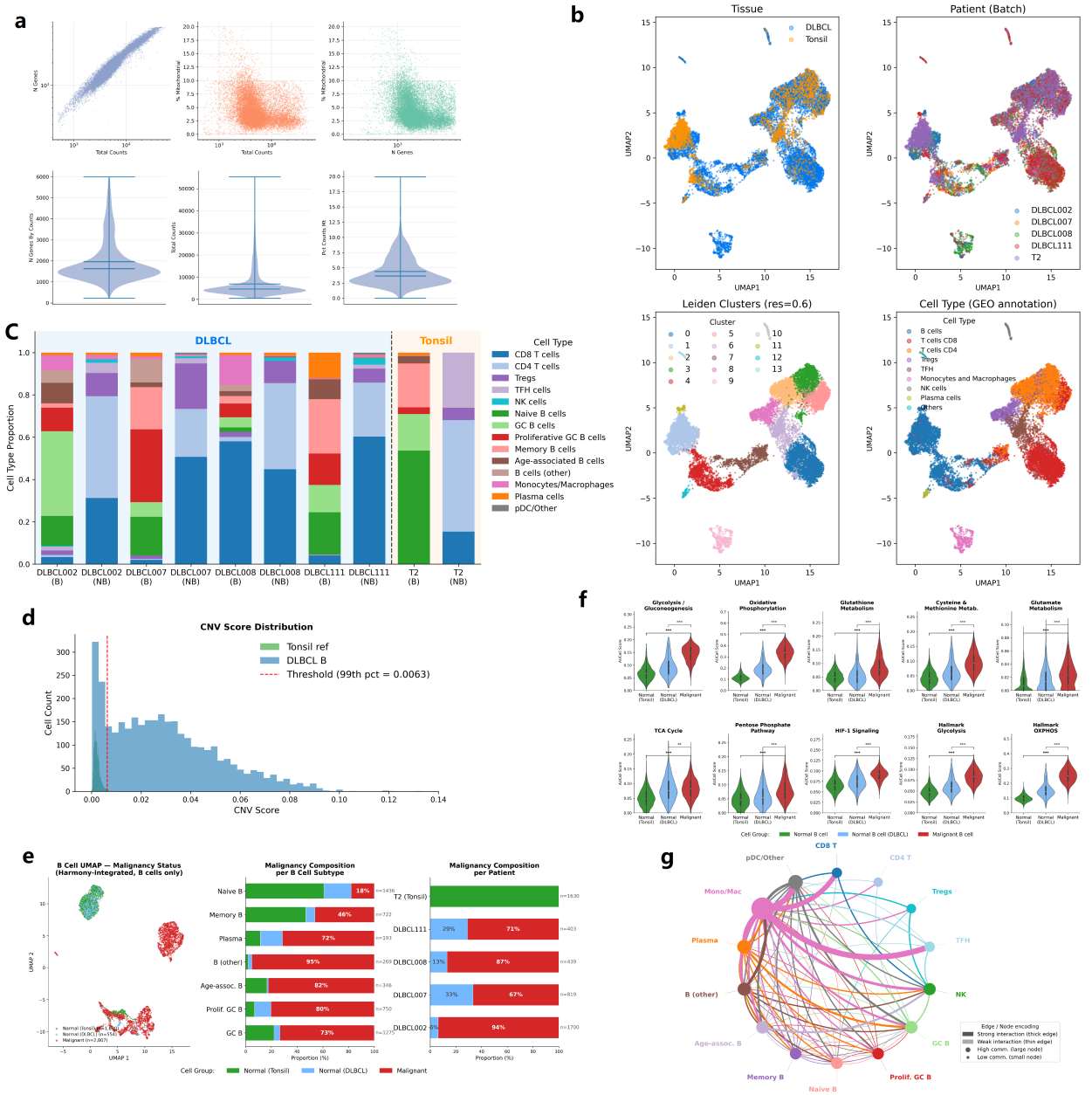

Figure S1: Data quality, dataset integration, malignant B-cell classification, metabolic activity, and communication structure in the DLBCL single-cell atlas. (a) Quality-control distributions of detected genes, total counts, and mitochondrial transcript percentage across retained single cells, shown as pairwise scatter plots (top) and violin plots (bottom) (b) Overview of the Harmony-integrated GSE182434 dataset showing tissue origin, patient/batch distribution, Leiden clusters, and original cell-type annotation (c) Per-sample cell-type composition across DLBCL and tonsil samples, highlighting inter-sample heterogeneity in both B-cell and non-B-cell compartments (d) Distribution of inferCNV scores in tonsil reference B cells and DLBCL B cells. The dashed line indicates the 99th percentile threshold of the tonsil reference distribution used to define inferCNV-high malignant B cells (e) B-cell-only UMAP showing malignancy assignment based on inferCNV scores (left), together with the proportion of malignant and normal-like cells across B-cell subtypes (middle) and across patients (right) (f) Violin plots showing AUCell score distributions for representative metabolic pathways across tonsil-derived normal B cells, inferCNV-low normal-like B cells from DLBCL samples, and inferCNV-high malignant B cells (g) Expanded cell–cell communication network across refined B-cell states and immune populations in DLBCL, inferred from top-ranked ligand–receptor interactions

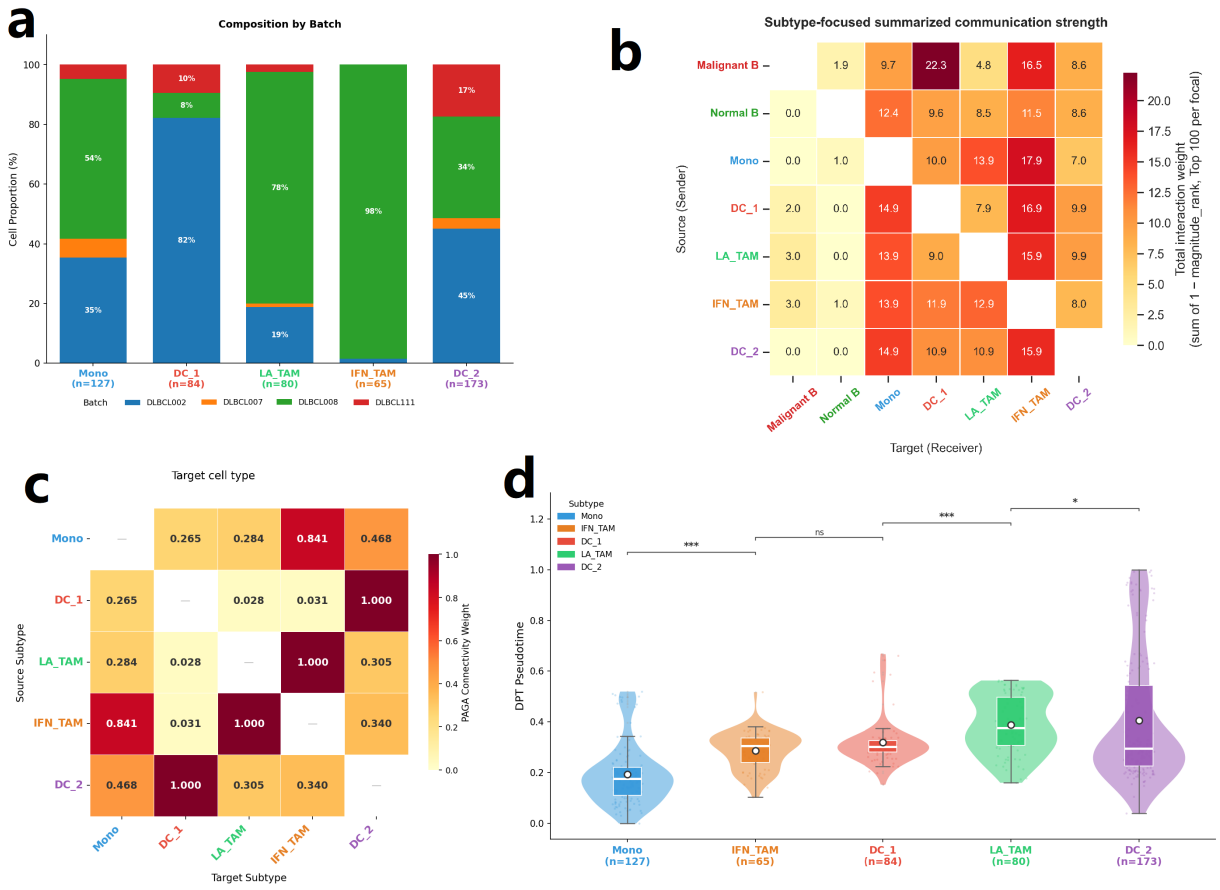

Figure S2: Myeloid-state heterogeneity, communication context, and branch-specific trajectory structure in DLBCL. (a) Sample-level composition of myeloid subtypes across DLBCL tumors, showing inter-patient heterogeneity in the relative abundance of Mono, DC 1, DC 2, IFN TAM, and LA TAM states (b) Subtype-focused summarized communication strength across malignant B cells, normal B cells, and myeloid subtypes, supporting the communication context shown in Fig 2c (c) PAGA connectivity matrix across myeloid subtypes, showing strongest branch connectivity between IFN TAM and LA TAM and substantial connectivity between Mono and IFN TAM along the TAM-oriented trajectory (d) Distribution of diffusion pseudotime (DPT) values across myeloid subtypes, supporting later pseudotime placement of LA TAM relative to Mono and intermediate positioning of IFN TAM

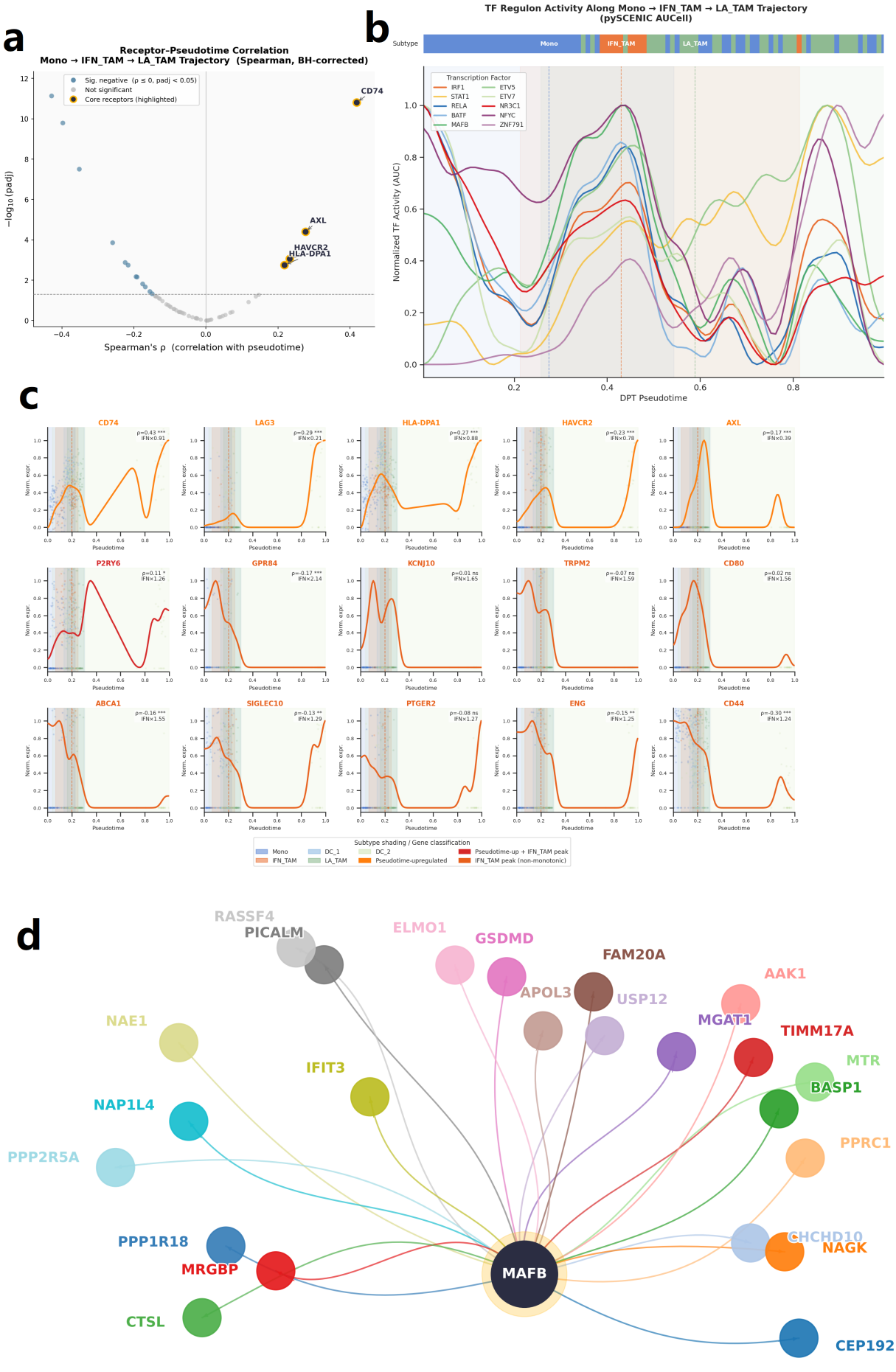
a
Receptor–Pseudotime Correlation
Mono → IFN_TAM → LA_TAM Trajectory (Spearman, BH-corrected)
Sig. negative (ρ ≤ 0, padj < 0.05)
Not significant
Core receptors (highlighted)
CD74
AXL
HAVCR2
HLA-DPA1
−log10 (padj)
Spearman's ρ (correlation with pseudotime)
b
TF Regulon Activity Along Mono → IFN_TAM → LA_TAM Trajectory
(pySCENIC AUCell)
Subtype
Mono
IFN_TAM
LA_TAM
Transcription Factor
IRF1
STAT1
RELA
BATF
MAFB
ETV5
ETV7
NR3C1
NFYC
ZNF791
Normalized TF Activity (AUC)
DPT Pseudotime
c
CD74
LAG3
HLA-DPA1
HAVCR2
AXL
P2RY6
GPR84
KCNJ10
TRPM2
CD80
ABCA1
SIGLEC10
PTGER2
ENG
CD44
Norm. expr.
Pseudotime
Subtype shading / Gene classification
Mono
IFN_TAM
DC_1
LA_TAM
DC_2
Pseudotime-upregulated
Pseudotime-up + IFN_TAM peak
IFN_TAM peak (non-monotonic)
d
RASSF4
PICALM
ELMO1
GSDMD
FAM20A
AAK1
APOL3
USP12
MGAT1
TIMM17A
NAE1
IFIT3
MTR
BASP1
NAP1L4
PPP2R5A
PPRC1
PPP1R18
MRGBP
CHCHD10
NAGK
MAFB
CTSL
CEP192

Figure S3: Receptor prioritization, regulon dynamics, and *MAFB*-centered regulatory architecture along the TAM-oriented trajectory. (a) Spearman correlation analysis between receptor expression and pseudotime across the Mono–IFN TAM–LA TAM trajectory. Highlighted points indicate receptors with significant positive correlation and prioritized relevance for downstream analysis (b) Smoothed regulon-activity curves of representative transcription factors across pseudotime, showing dynamic changes in TF activity along the Mono–IFN TAM–LA TAM trajectory (c) Smoothed pseudotime-expression curves for additional receptors derived from malignant B cell–enriched ligand–receptor interactions, illustrating heterogeneous temporal patterns across the TAM-oriented trajectory (d) Network representation of the motif-pruned *MAFB* regulon, showing predicted target genes connected to *MAFB* based on RcisTargetsupported motif enrichment.

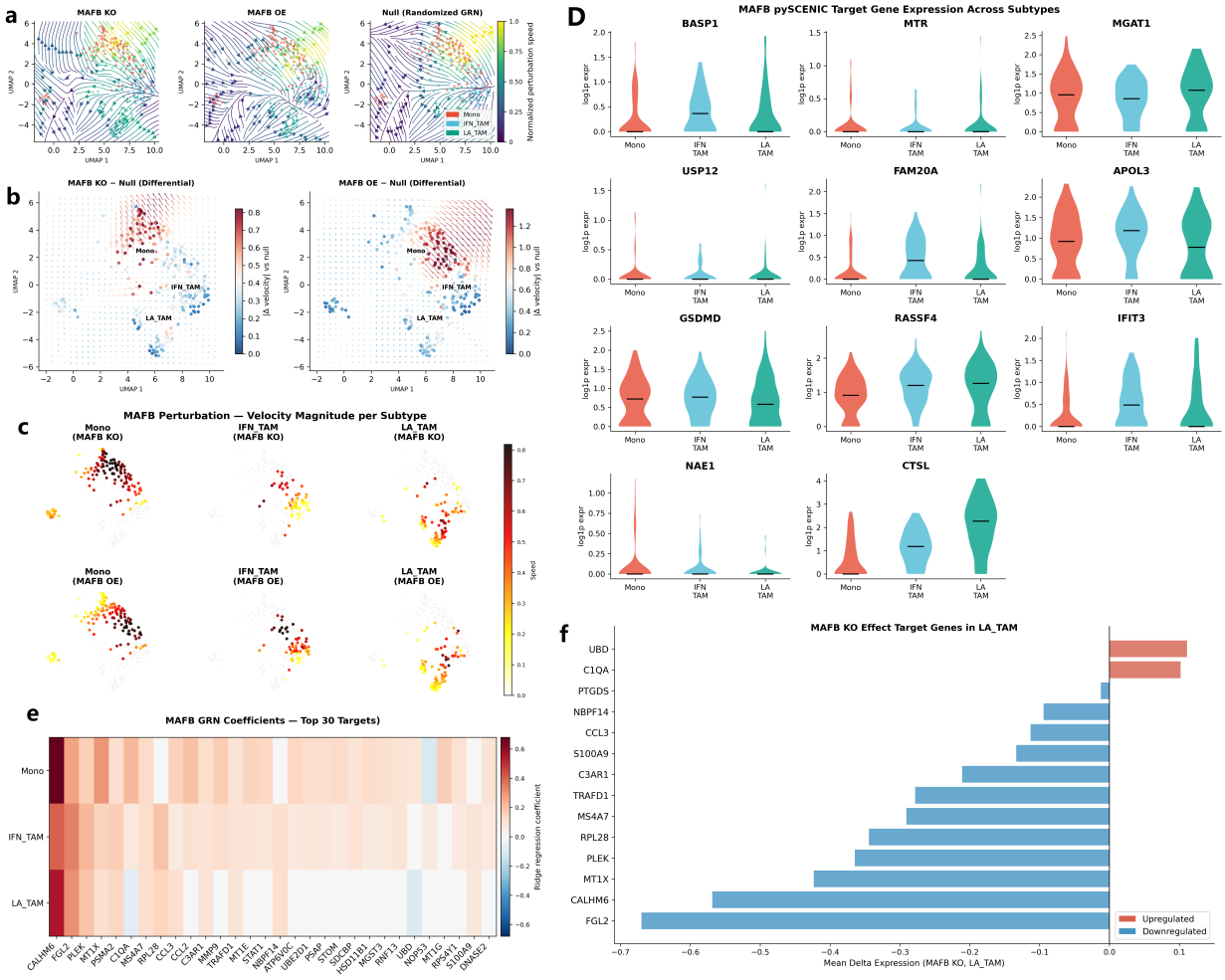

Figure S4: CellOracle-based *MAFB* perturbation effects on transition fields, subtype-specific perturbation magnitude, and downstream target responses. (a) Streamline visualization of CellOracle-inferred transition fields under simulated *MAFB* knockout and overexpression conditions, together with the randomized-network null control (b) Differential vector-field maps comparing simulated *MAFB* perturbation conditions with the randomized-network null control (c) Subtype-specific distributions of perturbation magnitude following simulated *MAFB* knockout and overexpression (d) Expression distributions of representative motif-pruned *MAFB* target genes across Mono, IFN TAM, and LA TAM states (e) Heatmap of inferred *MAFB* gene regulatory network coefficients for representative downstream targets across macrophage subtypes (f) Differential effects of simulated *MAFB* knockout on representative downstream target genes within LA TAM cells

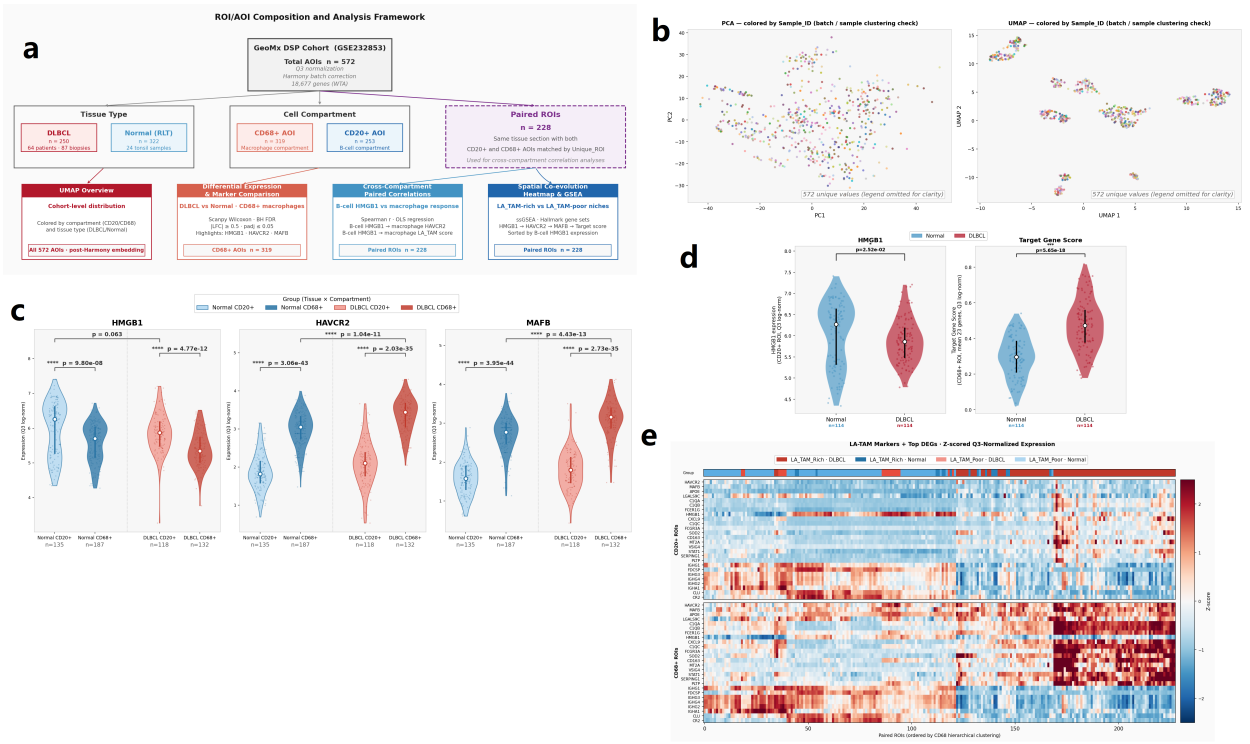

Figure S5: Spatial compartmentalization, cross-compartment coupling, and niche-associated pathway states in the GeoMx DSP cohort. (a) Schematic overview of the GeoMx DSP cohort composition and analysis framework, including AOI selection, compartment definitions, paired ROI matching, and downstream analytical modules (b) PCA and UMAP projections colored by sample identifier, showing the absence of dominant batch-driven clustering across AOIs after preprocessing (c) Compartment-stratified violin plots comparing *HMGB1*, *HAVCR2*, and *MAFB* expression across normal and DLBCL CD20$^+$ and CD68$^+$ AOIs (d)Violin plots comparing *HMGB1* expression (left) and target gene score (right) between normal and DLBCL CD20+ compartments, showing reduced *HMGB1* expression but increased target gene program activity in DLBCL-associated B-cell regions. (e) Paired-ROI heatmap showing coordinated variation in B-cell *HMGB1*, macrophage *HAVCR2*, macrophage LA TAM score, and target-gene program scores across matched DLBCL and normal regions

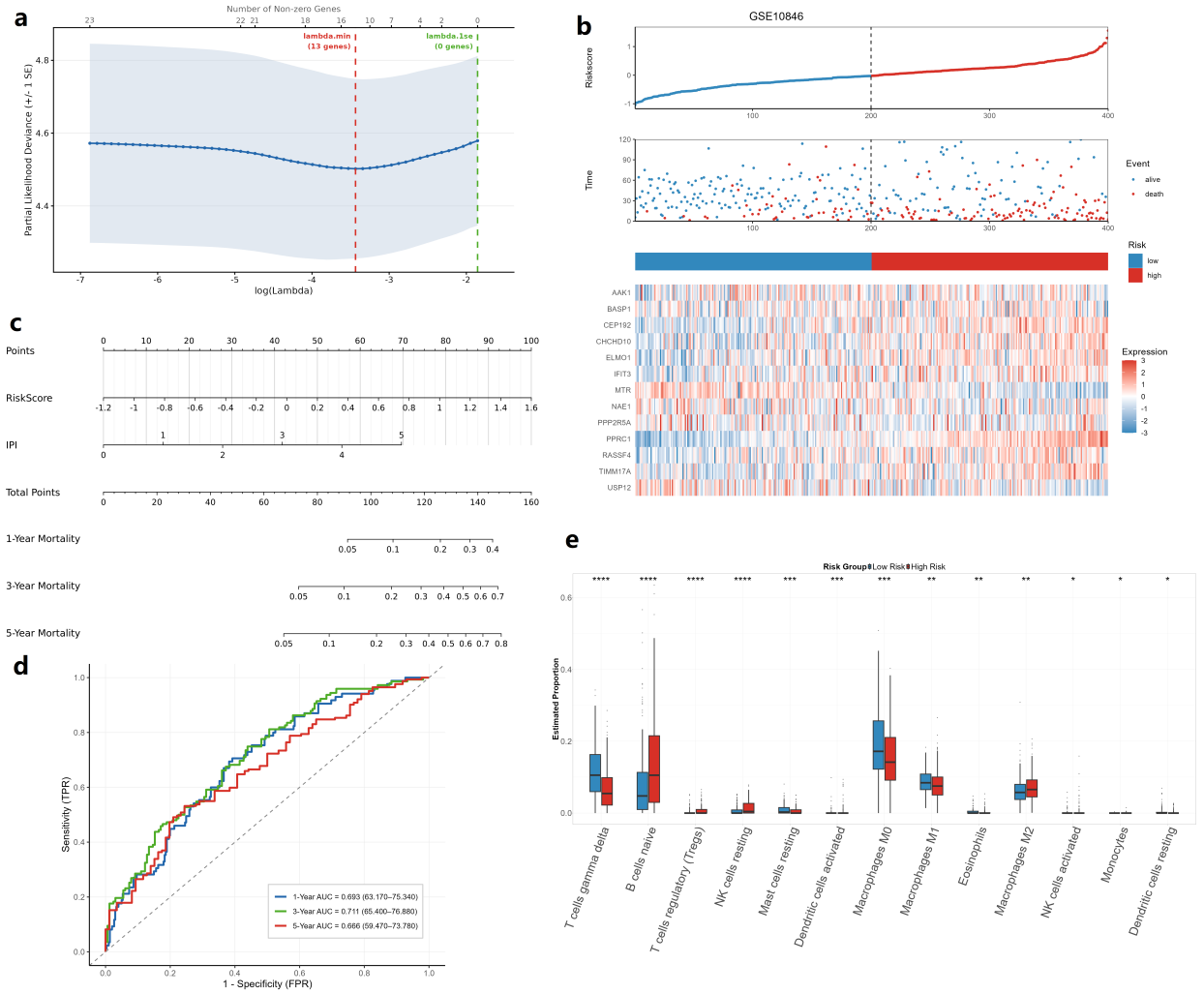

Figure S6: Construction, validation, and clinical interpretation of the MAFB regulon-based prognostic model. (a) Cross-validation profile used to determine the optimal penalty parameter in LASSO-Cox regression of motif-pruned MAFB regulon genes (b) Patient-level risk score distribution, survival status, and expression heatmap of the 13-gene prognostic model in the GSE10846 training cohort (c) Nomogram integrating the MAFB regulon-based risk score and IPI for prediction of 1-, 3-, and 5-year overall survival (d) Time-dependent ROC curves evaluating predictive performance of the prognostic model in the training cohort (e) Additional immune deconvolution comparisons between low-risk and high-risk tumors across inferred immune-cell subsets

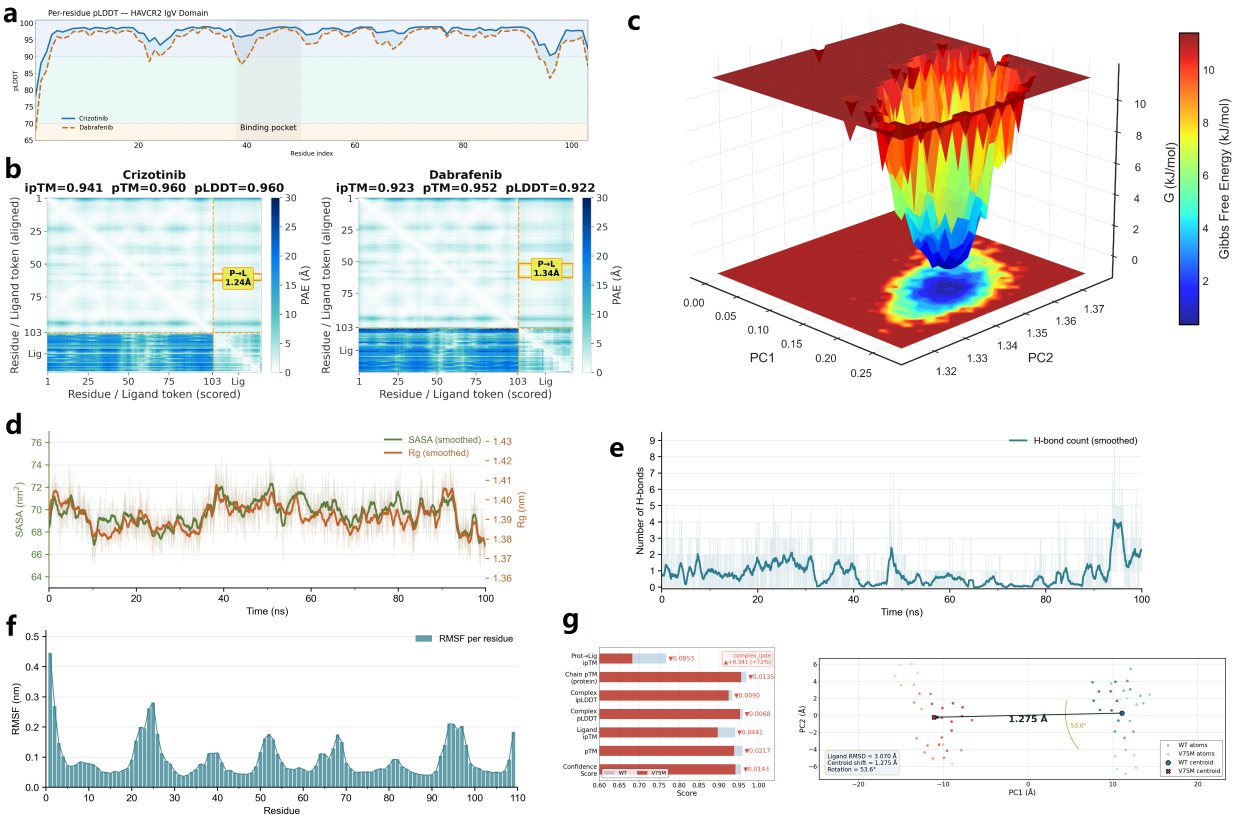

Figure S7: Supplementary structural confidence and molecular dynamics analyses supporting HAVCR2–Crizotinib prioritization and V75M-associated perturbation. (a) Per-residue confidence profiles for the HAVCR2 complexes with Crizotinib and dabrafenib, with the candidate binding-pocket interval highlighted (b) Pocket-focused protein-to-ligand PAE maps for the HAVCR2 complexes with Crizotinib and dabrafenib, supporting more favorable structural compatibility for Crizotinib than for dabrafenib across the prioritized pocket region (c) Three-dimensional free-energy landscape (FEL) derived from the HAVCR2–Crizotinib molecular dynamics trajectory, showing a dominant low-energy basin consistent with a preferred conformational state (d) Combined solvent-accessible surface area (SASA) and radius of gyration (Rg) profiles for the HAVCR2–Crizotinib complex over the 100-ns trajectory, supporting preservation of overall compactness (e) Time-dependent hydrogen-bond counts for the HAVCR2–Crizotinib complex during molecular dynamics simulation (f) Residue-level RMSF profile of HAVCR2 from the molecular dynamics simulation, showing localized flexibility across the modeled ectodomain (g) Additional structural comparison between the wild-type and V75M HAVCR2–Crizotinib complexes. Left, summary of confidence-related metrics for the wild-type and V75M models, including ligand ipTM, chain ipTM, complex ipLDDT, pocket ipLDDT, confidence score, and the corresponding differences (V75M minus wild type). Right, PCA-based geometric comparison showing mutation-associated ligand-pose divergence, including ligand RMSD, centroid shift, and rotation angle