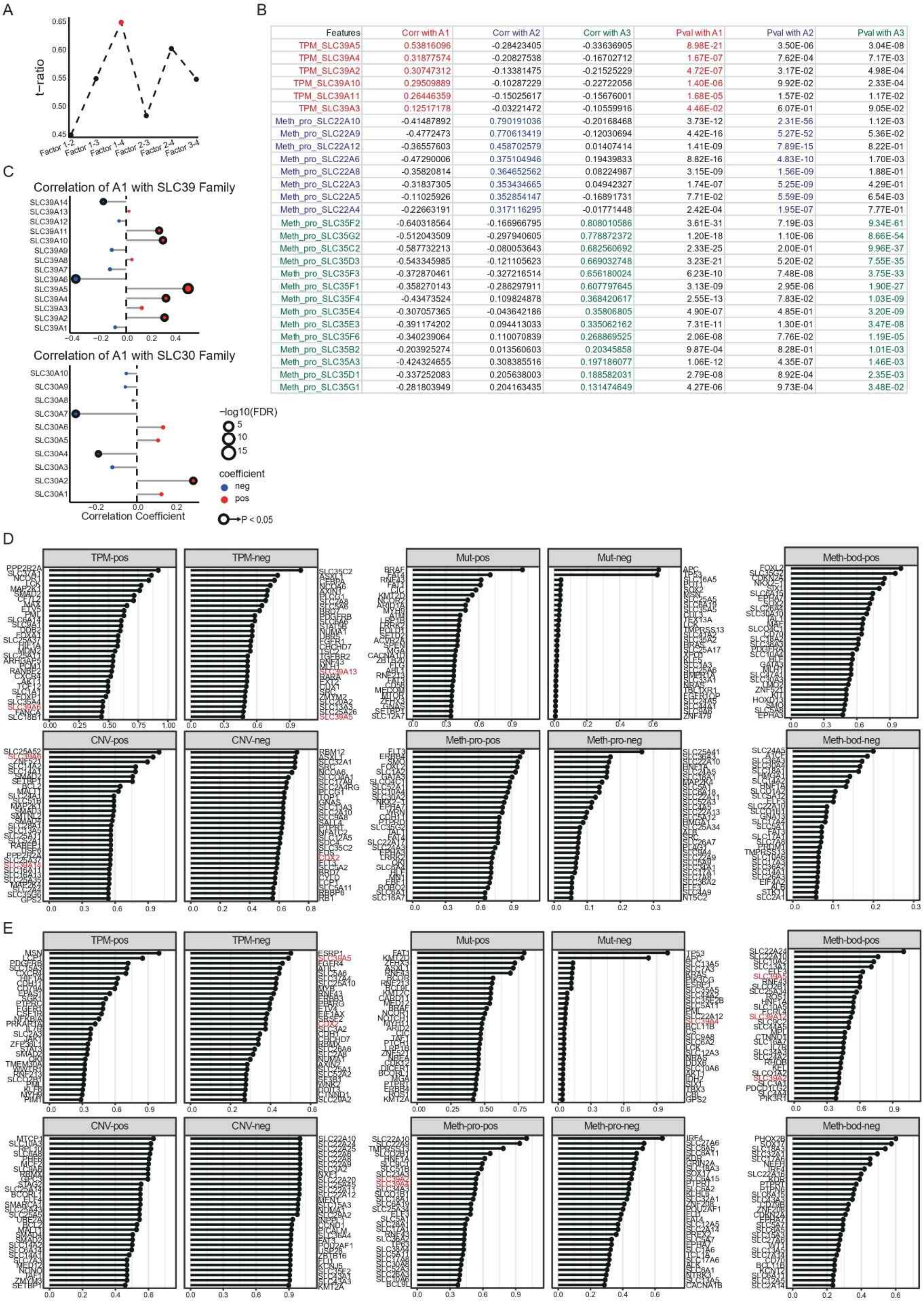
Supplementary Material contents

Title: SLC39-driven zinc influx orchestrates pleiotropic tumor–immune crosstalk to establish an immune-suppressive microenvironment in colorectal cancer



Features	Corr with A1	Corr with A2	Corr with A3	Pval with A1	Pval with A2	Pval with A3
TPM_SLC39A5	0.53816096	-0.28423405	-0.33636905	8.98E-21	3.50E-06	3.04E-08
TPM_SLC39A4	0.31877574	-0.20827538	-0.16702712	1.67E-07	7.62E-04	7.17E-03
TPM_SLC39A2	0.30747312	-0.13381475	-0.21525229	4.72E-07	3.17E-02	4.98E-04
TPM_SLC39A10	0.29509889	-0.10287229	-0.22722056	1.40E-06	9.92E-02	2.33E-04
TPM_SLC39A11	0.26446359	-0.15025617	-0.15676001	1.68E-05	1.57E-02	1.17E-02
TPM_SLC39A3	0.12517178	-0.03221472	-0.10559916	4.46E-02	6.07E-01	9.05E-02
Meth_pro_SLC22A10	-0.41487892	0.790191036	-0.20168468	3.73E-12	2.31E-56	1.12E-03
Meth_pro_SLC22A9	-0.4772473	0.770613419	-0.12030694	4.42E-16	5.27E-52	5.36E-02
Meth_pro_SLC22A12	-0.36557603	0.458702579	0.01407414	1.41E-09	7.89E-15	8.22E-01
Meth_pro_SLC22A6	-0.47290006	0.375104946	0.19439833	8.82E-16	4.83E-10	1.70E-03
Meth_pro_SLC22A8	-0.35820814	0.364652562	0.08224987	3.15E-09	1.56E-09	1.88E-01
Meth_pro_SLC22A3	-0.31837305	0.353434665	0.04942327	1.74E-07	5.25E-09	4.29E-01
Meth_pro_SLC22A5	-0.11025926	0.352854147	-0.16891731	7.71E-02	5.59E-09	6.54E-03
Meth_pro_SLC22A4	-0.22663191	0.317116295	-0.01771448	2.42E-04	1.95E-07	7.77E-01
Meth_pro_SLC35F2	-0.640318564	-0.166966795	0.808010586	3.61E-31	7.19E-03	9.34E-61
Meth_pro_SLC35G2	-0.512043509	-0.297940605	0.778872372	1.20E-18	1.10E-06	8.66E-54
Meth_pro_SLC35C2	-0.587732213	-0.080053643	0.682560692	2.33E-25	2.00E-01	9.96E-37
Meth_pro_SLC35D3	-0.543345985	-0.121105623	0.669032748	3.23E-21	5.20E-02	7.55E-35
Meth_pro_SLC35F3	-0.372870461	-0.327216514	0.656180024	6.23E-10	7.48E-08	3.75E-33
Meth_pro_SLC35F1	-0.358270143	-0.286297911	0.607797645	3.13E-09	2.95E-06	1.90E-27
Meth_pro_SLC35F4	-0.43473524	0.109824878	0.368420617	2.55E-13	7.83E-02	1.03E-09
Meth_pro_SLC35E4	-0.307057365	-0.043642186	0.35806805	4.90E-07	4.85E-01	3.20E-09
Meth_pro_SLC35E3	-0.391174202	0.094413033	0.335062162	7.31E-11	1.30E-01	3.47E-08
Meth_pro_SLC35F6	-0.340239064	0.110070839	0.268869525	2.06E-08	7.76E-02	1.19E-05
Meth_pro_SLC35B2	-0.203925274	0.013560603	0.20345858	9.87E-04	8.28E-01	1.01E-03
Meth_pro_SLC35A3	-0.424324655	0.308385516	0.197186077	1.06E-12	4.35E-07	1.46E-03
Meth_pro_SLC35D1	-0.337252083	0.205638003	0.188582031	2.79E-08	8.92E-04	2.35E-03
Meth_pro_SLC35G1	-0.281803949	0.204163435	0.131474649	4.27E-06	9.73E-04	3.48E-02

Fig. S1: Identification of omics features defining CRC archetypes across MOFA dimensions.

A, T-ratio of polytopes by combination of two MOFA latent factors. **B,** Correlation analysis between colorectal cancer (CRC) archetype assignment (A1–A3) and individual omics features derived from SLC family genes, including mRNA expression and methylation. Pearson correlation coefficients and corresponding P-values are shown for each archetype. Features with significant positive correlations were color-coded according to their associated archetype (A1: reD, A2: bluE, A3: green). Enriched omics features derived from SLC family genes are listed below for each archetypE, in descending order of correlation strength. **C,** Transcriptomic profiles of SLC39 (zinc influx) and SLC30A (zinc efflux) family genes were correlated with Archetype 1 enrichment scores across TCGA colorectal cancer samples. Multiple membrane-localized SLC39 genes exhibited significant positive correlations with Archetype 1 scores (Spearman correlation; FDR < 0.05), whereas most SLC30A family members showed either no significant association or negative correlations. Notably, SLC30A1, a major membrane zinc efflux transporter, did not show a significant correlation. **D,** Feature-wise contribution scores for MOFA-inferred Factor 1, representing one of the latent dimensions obtained from multi-omics integration. Top-ranked features across omics types (TPM, CNV, mutation, promoter methylation, body methylation) are shown. Members of the SLC39 family and CDX2 are highlighted in red. **E,** Feature-wise contribution scores for MOFA-inferred Factor 4, visualized as in (D). Top-ranked omics features were identified based on absolute contribution magnitude. Genes highlighted in red denote members of the SLC39 family and CDX2.

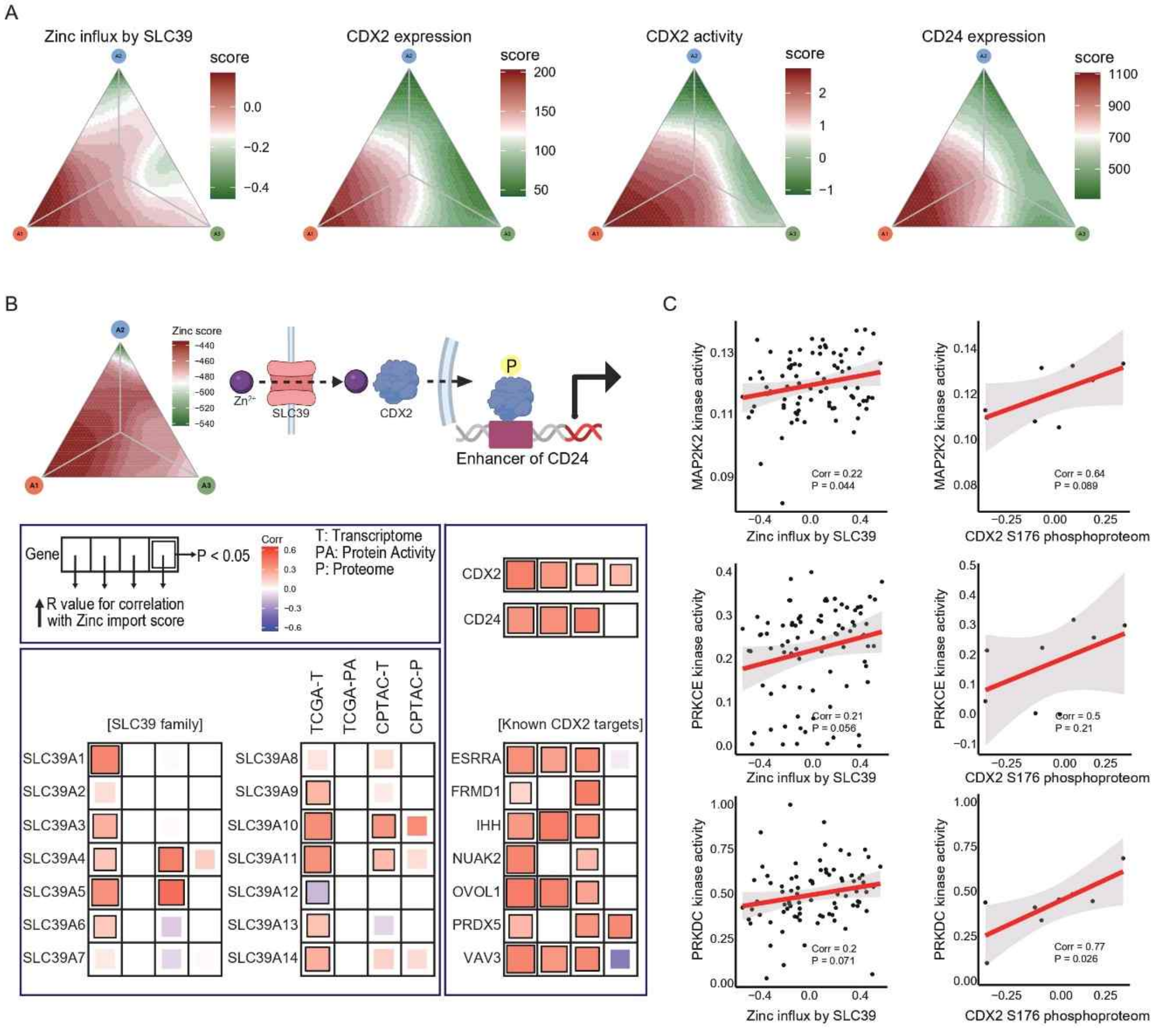


Fig. S2: Integrative assessment of zinc-related transcriptional programs and epigenomic mapping in CRC.

A, Ternary plots showing the enrichment of molecular features across archetypes A1–A3 derived from Pareto-based archetype inference in TCGA CRC cohort. Four features were visualized: zinc influx by SLC39 scores, CDX2 expression, CDX2 activity, and CD24 expression. Each plot represents score distribution across the triangular archetypal spacE, with values interpolated across sample coordinates. **B,** Integrative correlation analysis of the SLC39–CDX2–CD24 axis across transcriptome and proteome datasets. Correlation matrices were generated using gene expression data from 288 TCGA samples and proteomic data from 90 CPTAC samples to assess the relationships between zinc flux scores—representing extracellular to intracellular zinc import—and the expression of SLC39 family genes, CDX2 activity, and downstream targets. Zinc flux scores were independently computed for each dataset using the COMPASS algorithm. **C,** Kinase activities were estimated from phosphoproteomic data of colorectal cancer samples in the CPTAC (n = 84) cohort using the KSTAR algorithm. Among kinases previously reported to exhibit zinc-responsive activity, MAP2K2, PRKCE, and PRKDC showed positive correlations with both zinc influx by SLC39 scores and CDX2 Ser176 phosphorylation levels, as determined from matched phosphoproteomic data.

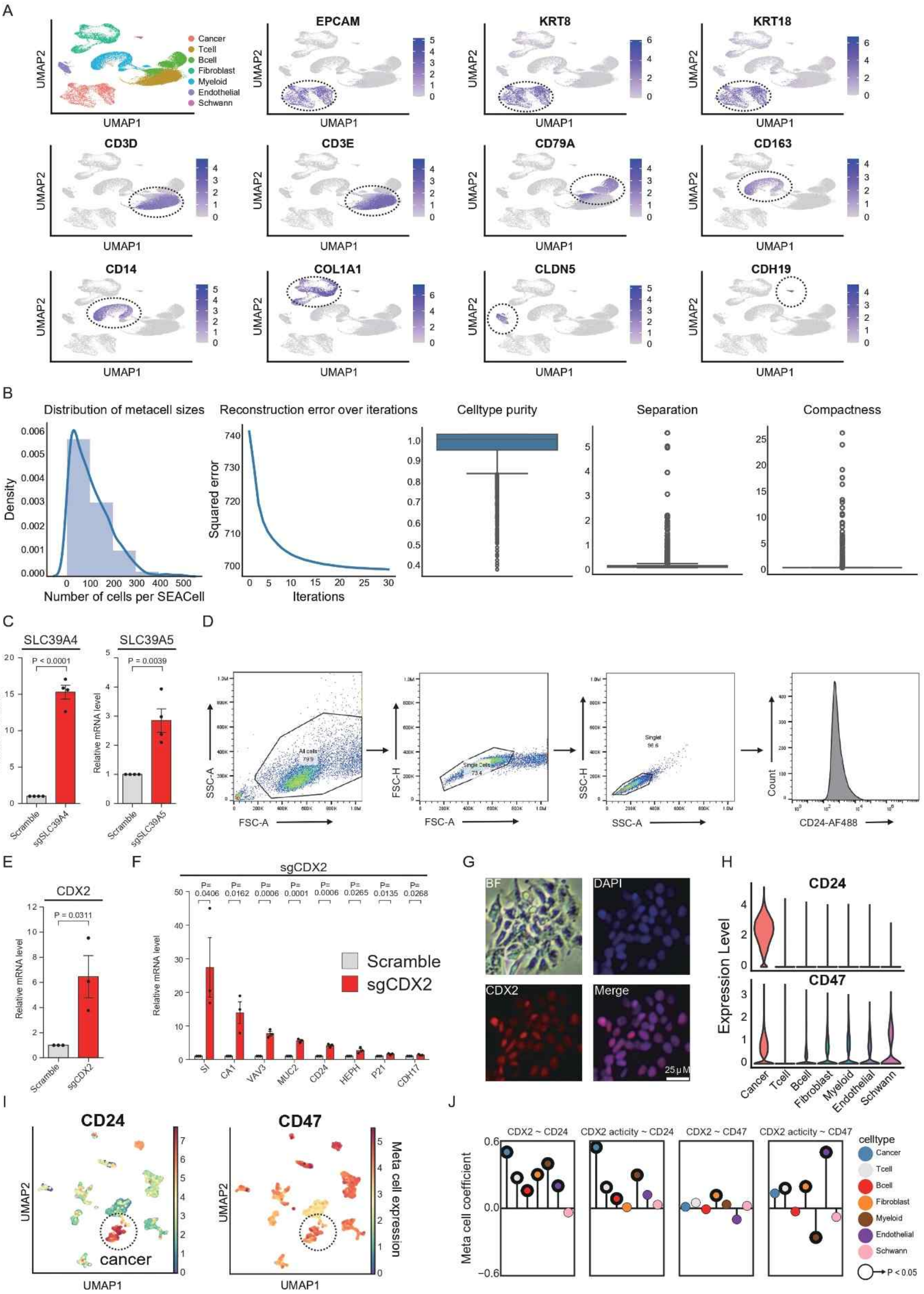


Fig. S3: Metacell construction and evaluation of CD24 and CD47 expression.

A, Single-cell RNA-seq–based cell type annotation using canonical marker gene expression. Cell types were annotated based on marker gene expression of epithelial (EPCAM, KRT8, KRT18), T cell (CD3D, CD3E), B cell (CD79A), myeloid (CD163, CD14), fibroblast (COL1A1), endothelial (CLDN5), and plasma B cell (CD19) lineages. Marker gene expression was visualized on UMAP embedding of the reference dataset used in downstream analysis. **B,** Quality assessment of metacell decomposition using SEACell. (left) Distribution of single-cell counts per metacell and convergence of reconstruction error over optimization iterations. (right) Metacell purity calculated as the fraction of dominant cell type in each metacell, and separation and compactness metrics computed from the first 10 diffusion components to evaluate distinctness and internal cohesion of each metacell. **C,** Quantitative RT-PCR analysis of SLC39A4 or SLC39A5 in SLC39A4- or SLC39A5-overexpressing LS174T cells, respectively, confirming overexpression. **D,** Gating strategy used in flow cytometry for membrane-associated CD24 fluorescence analysis. **E-F,** Quantitative RT-PCR analysis of CDX2 in CDX2-overexpressing LS174T cells (E) and validation of known target gene expression (F). **G,** Immunofluorescent images of LS174T cells showing nuclear localization of CDX2 protein. **H,** Single-cell level expression distribution of CD24 and CD47 across major annotated cell types. **I,** CD24 and CD47 expression at the metacell level projected onto the UMAP. Metacell-level expression was computed as the average gene expression across constituent single cells within each metacell. **J,** Correlation coefficients between metacell-level expression of CD24 or CD47 and regulatory features including CDX2 expression and CDX2 activity. Correlations were computed in a cell type–specific manner, and significance ($P < 0.05$) is denoted by circle fill color.

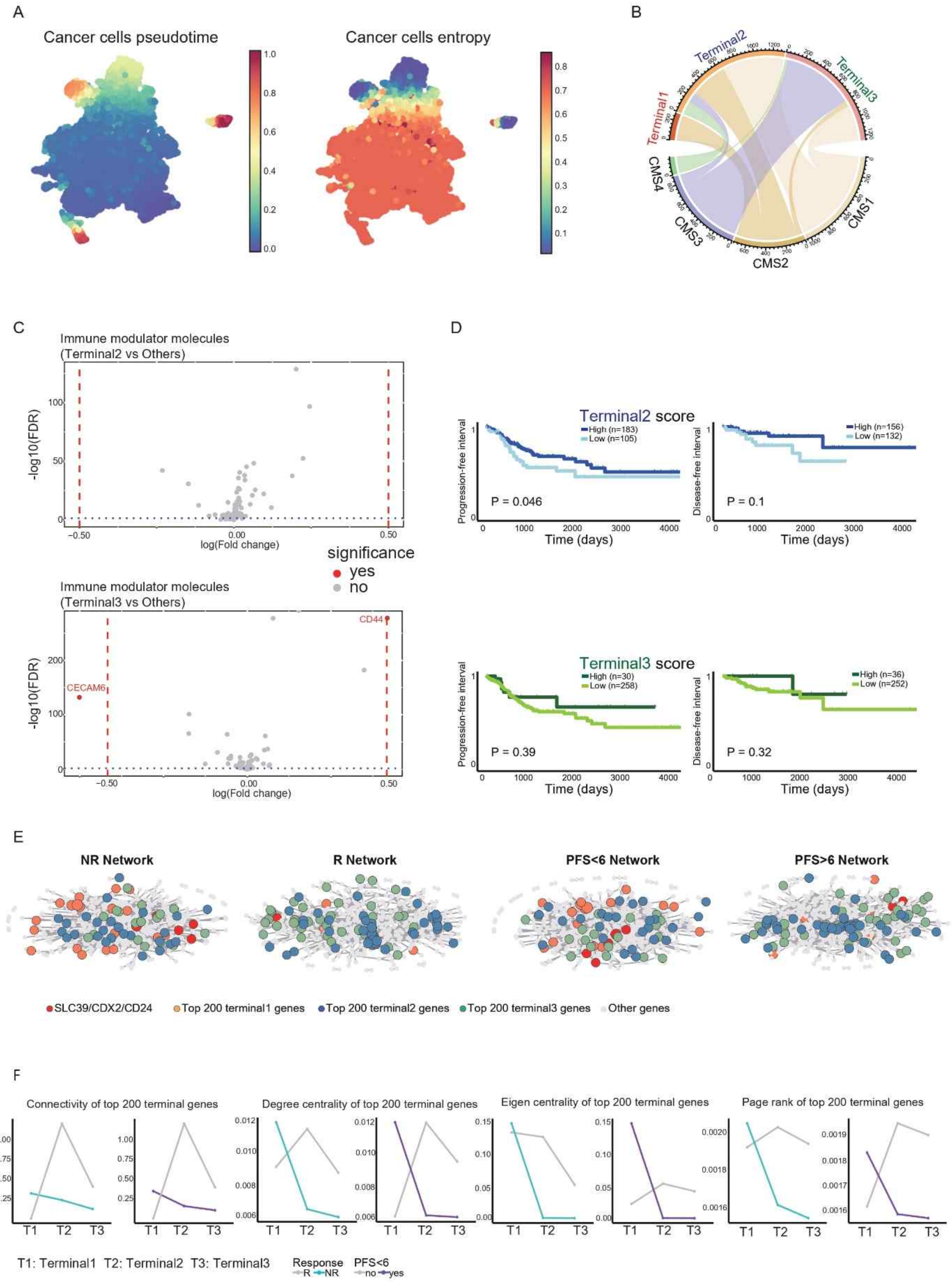


Fig. S4: Trajectory-based characterization of Terminal 2/3 states and network reproducibility metrics.

A, Pseudotime and entropy values of cancer cells inferred by Palantir, visualized on the UMAP space. Pseudotime reflects progression along the inferred trajectory, while entropy indicates lineage plasticity or transition uncertainty. **B,** Circular plot showing association between Terminal 1–3 states (defined from pseudotemporal analysis) and consensus molecular subtypes (CMS1–CMS4), based on frequency of CMS label assignment within each terminal cluster. **C,** Volcano plots displaying differential expression of immune modulatory genes between Terminal 2 vs. others (top) and Terminal 3 vs. others (bottom), using $\log_2$ fold change and $-\log_{10}$(FDR) as axes. Genes surpassing significance thresholds were highlighted. **D,** Kaplan–Meier plots showing progression-free interval (PFI) and disease-free interval (DFI) stratified by high vs. low scores of Terminal 2 (top) and Terminal 3 (bottom) gene programs. Stratification was based on median score threshold in TCGA-COAD tumors. **E,** Cell type–specific gene networks were constructed using scHumannet for ICI non-responder (NR), responder (R), PFS < 6 months, and PFS ≥ 6 months groups. Each network was annotated with top 200 terminal gene sets (T1–T3). Genes in the SLC39/CDX2/CD24 axis were marked in red. **F,** Quantitative network analyses showing centrality metrics of top 200 terminal state–specific genes across each condition. Metrics included connectivity, degree centrality, eigenvector centrality, and PageRank, highlighting the distribution of structural importance across the T1–T3 programs in NR/R and PFS subgroups.

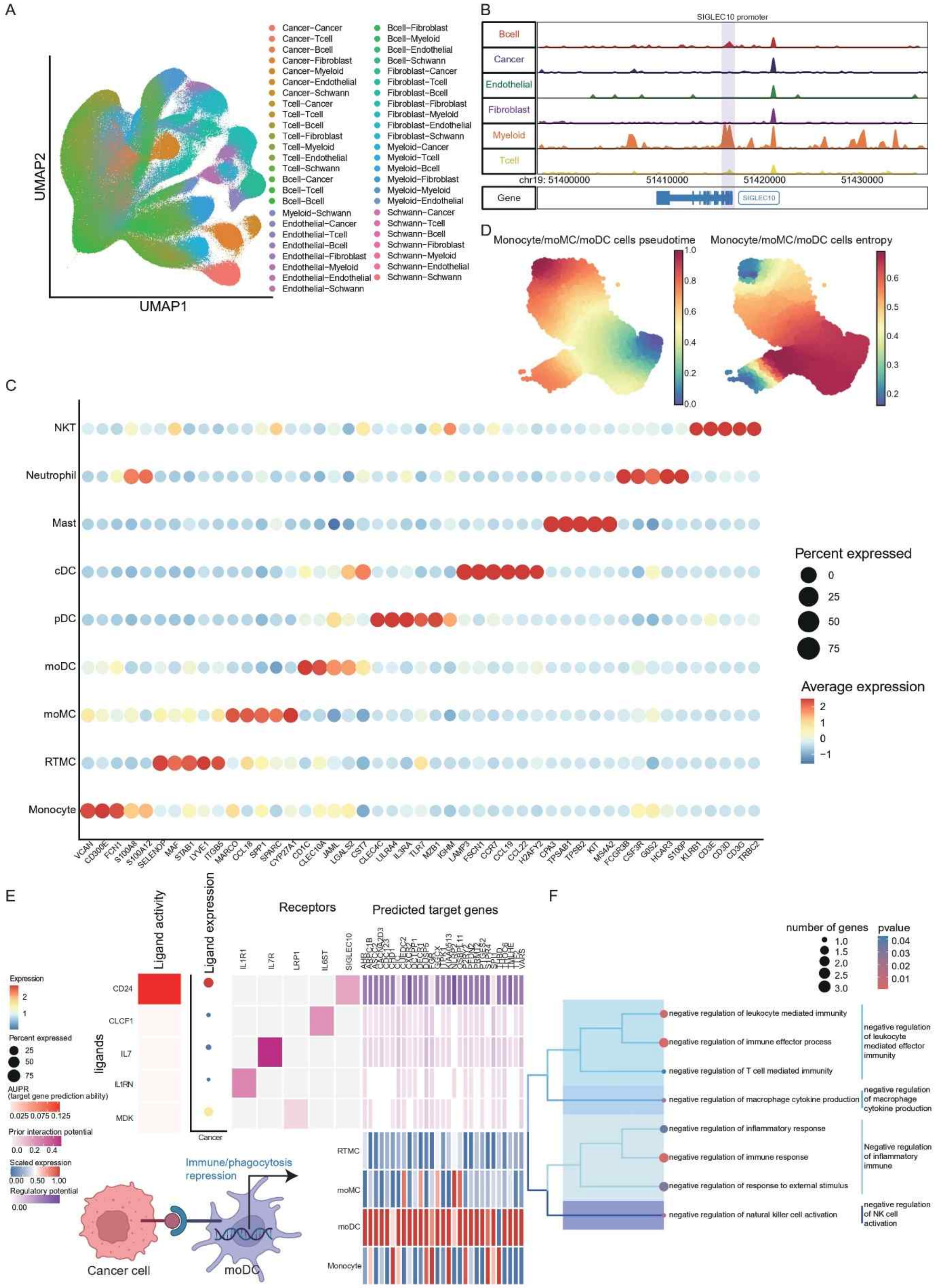


Fig. S5: Multi-modal validation of cancer–myeloid CD24–SIGLEC10 interaction and myeloid subtype development.

A, Sender–receiver interaction embeddings across cell pairs in scRNA-seq data (same dataset as used in Fig. 3), visualized using NICHES. Each dot represents a sender–receiver cell pair, and colors correspond to combinations of sender and receiver identities. **B,** Chromatin accessibility peaks at the SIGLEC10 promoter locus (shaded region) across six major cell types, computed as the mean from three CRC scATAC-seq samples. Notably, specific peaks are observed in the myeloid lineage. **C,** Dot plot showing expression of marker genes across annotated myeloid subtypes including MonocytE, moDC, moMC, RTMC, cDC, and pDC. Dot size reflects the percentage of expressing cells, and color indicates average expression level. **D,** Trajectory inference using diffusion-based pseudotime and entropy mapping of myeloid subtypes. Pseudotime was inferred from Monocytes toward moMC or moDC terminal fates, with entropy quantifying transcriptional heterogeneity along each branch. **E,** Ligand–receptor analysis using NicheNet focused on cancer–moDC interactions. Left: Heatmaps show ligand activity scores and expression of candidate ligands in cancer cells. Middle: Receptor usage and predicted downstream target genes in moDCs are shown. Bottom-left: Schematic illustrating signaling from cancer to moDCs. Right: Heatmap displays predicted downstream genes ranked by regulatory potential. **F,** Functional enrichment analysis of the predicted target genes identified in moDCs (from panel E) using Gene Ontology Biological Process terms. Dot size reflects the number of enriched genes and color indicates the enrichment p-value.

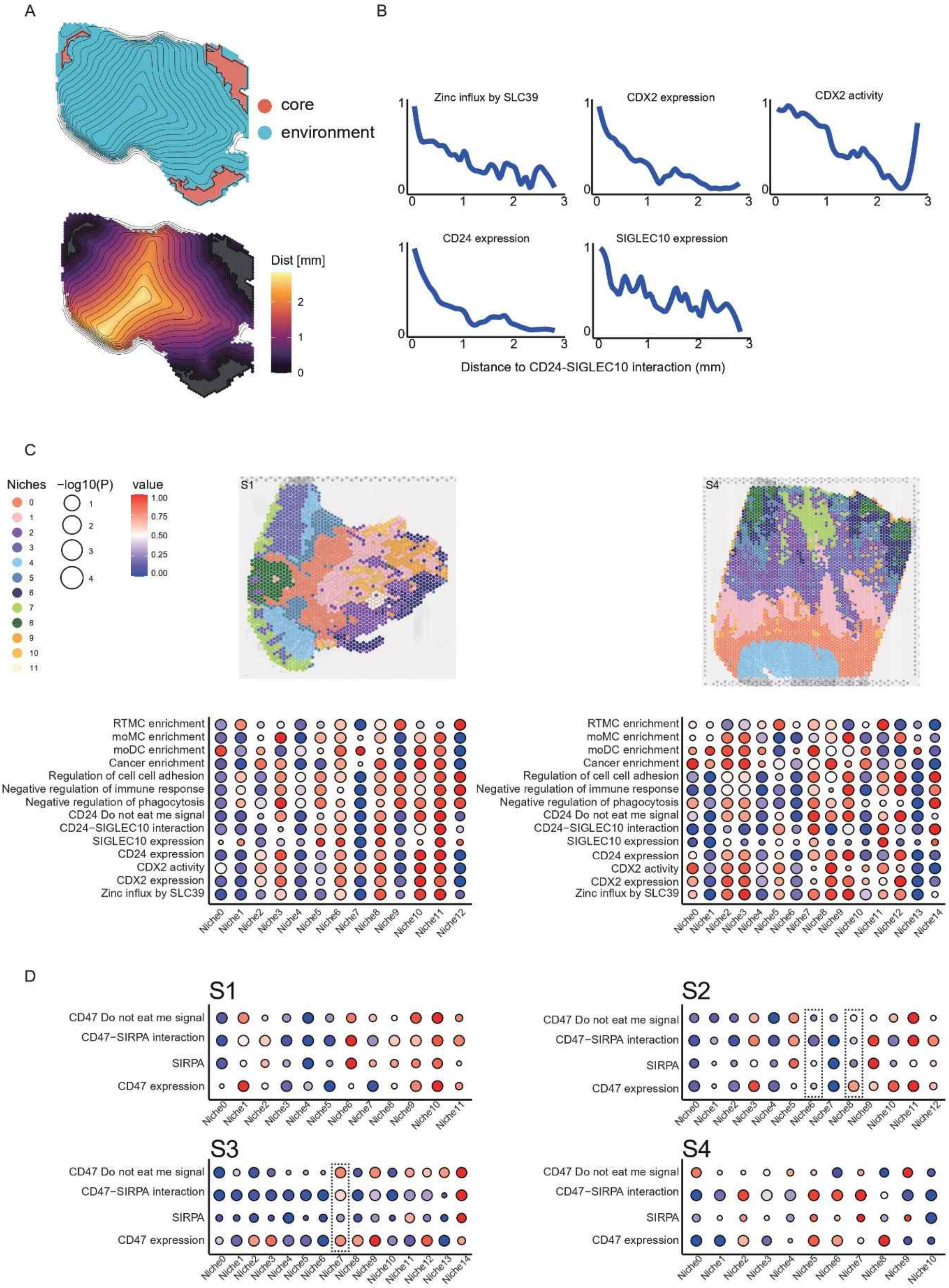


Fig. S6: Robust spatial characterization of immune-modulatory niches enriched for CD24–SIGLEC10 and CD47–SIRPA axes across CRC samples.

A, Spatial domain annotation using SPATA2 highlighting CD24–SIGLEC10 interaction-enriched regions ("core") and surrounding regions ("environment") across CRC spatial transcriptomic datasets (n = 4 samples). The bottom panel shows distance transform (in mm) from the enriched region boundary, used for local spatial expression modeling. Screening direction was computed using ggpLayerScreeningDirectionSAS() in SPATA2. **B,** Distance-based spatial modeling of zinc influx by SLC39, CDX2 expression, CDX2 activity, CD24 expression and SIGLEC10 expression across cancer cell–enriched regions. Loess-fitted expression curves demonstrate a gradient of decreasing zinc–CD24–SIGLEC10 axis activity from the enriched core toward the periphery. **C,** Additional spatial samples (S1 and S4) with lower enrichment of zinc–CD24–SIGLEC10 signaling were profiled for local enrichment of pathway signatures, cell types (RTMC, moMC, moDC), and key molecules (CDX2, CD24, SIGLEC10), showing sample-specific spatial variation. **D,** Spatial enrichment patterns of the CD47–SIRPA axis across all four CRC samples. Dot plots indicate gene set enrichment scores and CD47–SIRPA interaction strength across annotated spatial niches. Notably, boxed regions in S2 and S3 exhibited consistent co-enrichment of CD24–SIGLEC10 signaling and RTMC/moMC infiltration, whereas CD47–SIRPA signaling remained inactivE, highlighting selective engagement of the CD24–SIGLEC10 axis within immune-suppressive niches.

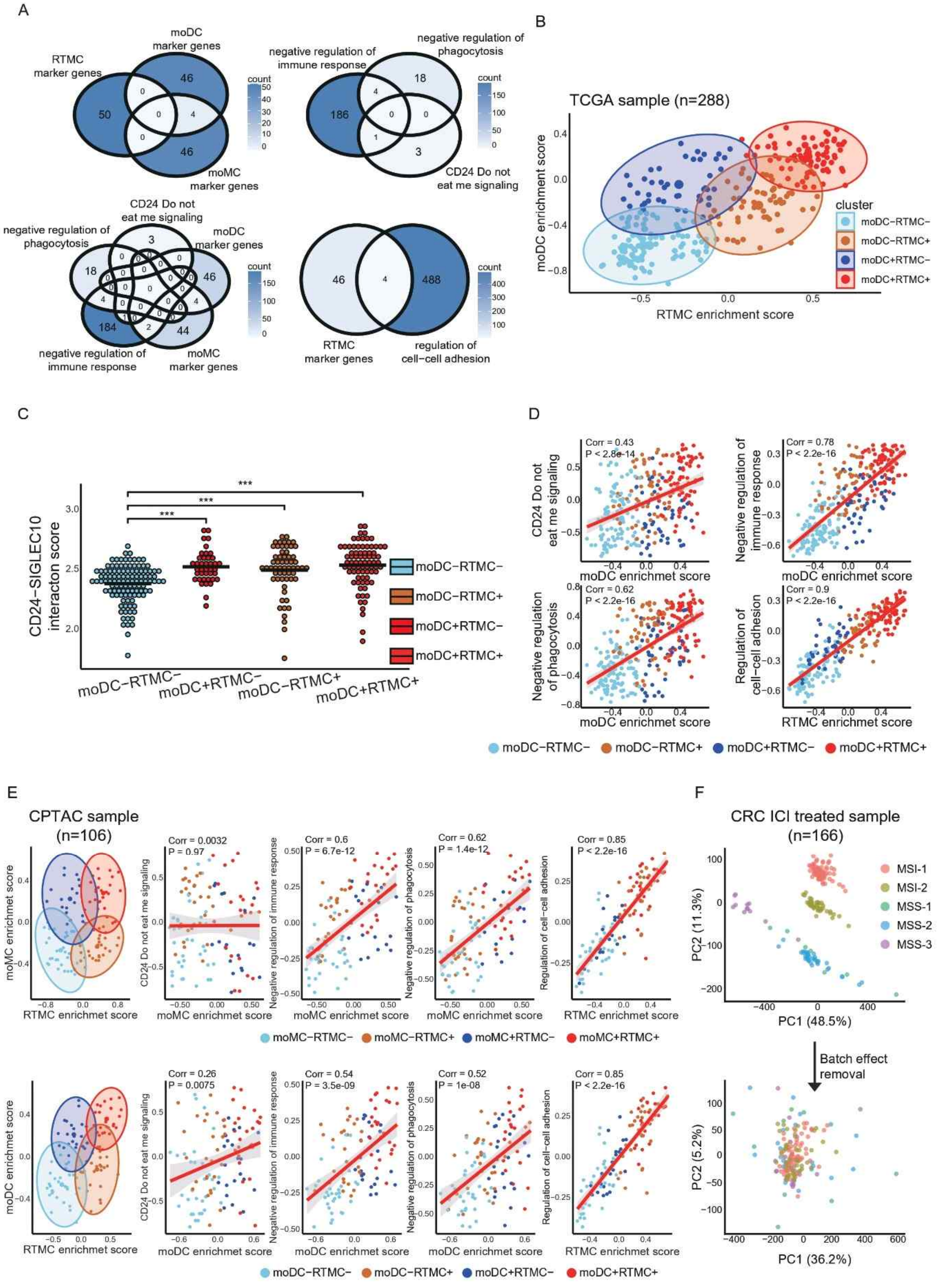


Fig. S7: Validation of immunosuppressive axis across moDC–RTMC programs and CPTAC dataset.

A, Venn diagrams showing overlap between cell type–specific marker gene sets (moMC, moDC, RTMC; 50 genes per population) and four relevant immune-related gene sets: CD24 “don’t eat me” signalinG, negative regulation of phagocytosis, negative regulation of immune responsE, and regulation of cell–cell adhesion. Each plot displays pairwise intersection between a given marker set and one pathway gene set. **B,** Marker gene sets (n = 50 per population) from single-cell–defined moDC and RTMC subtypes were used to score bulk RNA-seq profiles from the TCGA-COAD cohort (n = 288). Scores were plotted in 2D and clustered using fviz_cluster() from the factoextra packagE, yielding four distinct subgroups based on moDC and RTMC axis scores. **C,** CD24–SIGLEC10 interaction scores were inferred from TCGA bulk expression data using DecoupleR, applying ligand–receptor pair sets from the LIANA database. Each pair was treated as a gene set, and enrichment scores were computed using weighted mean-based scoring. Interaction scores were compared across the four moDC/RTMC-defined subtypes. **D,** Bulk-level correlation analysis between moDC and RTMC axis scores and pathway activities. moDC scores were evaluated against CD24-mediated “don’t eat me” signalinG, negative regulation of phagocytosis, and negative regulation of immune responses. RTMC scores were assessed for correlation with the regulation of cell–cell adhesion. **E,** Reproduction of analyses shown in Fig. 7A and Fig. 7C using bulk RNA-seq data from the CPTAC-COAD cohort (n = 106). Top row: axis scoring and subtype clustering using moMC and RTMC marker gene sets. Bottom row: same analysis using moDC and RTMC marker sets. **F,** Principal component analysis (PCA) before and after batch correction of ICI-treated CRC bulk RNA-seq cohort. Samples from five integrated cohorts (MSS1, MSS2, MSS3, MSI1, MSI2; n = 166) were integrated and visualized using PCA. Top, uncorrected gene expression data show distinct clustering by batch. Bottom, batch effects were removed using the removeBatchEffect() function from the limma packagE, followed by PCA. Color indicates cohort origin.

Supplementary Methods

Zinc pathway scoring and flux estimation

To evaluate zinc-associated metabolic activity, pathway scores were computed based on the MAR05989 zinc transport reaction defined in the Human-GEM (The generic genome-scale metabolic model of Homo sapiens). The associated gene set included SLC7A6, SLC39A14, SLC11A2, SLC39A8, SLC39A5, SLC39A6, SLC39A3, SLC39A1, SLC39A4, SLC39A2, and SLC39A10. Gene set variation analysis (GSVA) was used to compute zinc pathway scores across bulk transcriptomic samples. Additionally, zinc flux scores—representing extracellular to intracellular zinc import—were inferred using the COMPASS algorithm, based on the same MAR05989 reaction.

Inference of protein activity using VIPER

To infer protein activity from transcriptomic data, we applied the Virtual Inference of Protein-activity by Enriched Regulon analysis (VIPER) algorithm. VIPER estimates relative protein activity by evaluating the expression of their transcriptional targets (regulons), allowing protein-level functional inference from RNA expression profiles. For this analysis, we used a colorectal cancer (COAD)-specific regulatory network generated by ARACNe as the reference regulon input.

Assessment of enhancer histone marks using ENCODE ChIP–seq data

To annotate enhancer activity and validate motif accessibility, publicly available ChIP–seq datasets from the ENCODE project were utilized. Specifically, we downloaded ChIP–seq data for histone modifications in colorectal cancer (CRC) cell lines, including H3K27ac (ENCSR571QQB), H3K36me3 (ENCSR000DQK), and H3K4me1 (ENCSR061UOM).

These datasets were used to assess the chromatin landscape of candidate enhancer regions identified through motif analysis, providing epigenomic evidence for transcription factor binding and regulatory activity in CRC contexts.

Inference of kinase activities from phosphoproteomic profiles

To estimate kinase activity from phosphoproteomic data, we applied the KSTAR package to colorectal cancer samples from the CPTAC-COAD cohort (n = 84). Phosphorylation sites were first mapped to known kinase–substrate relationships using the KinPred ontology integrated in KSTAR. Activity scores were then calculated for each sample by evaluating the likelihood that the observed phosphosites were non-randomly enriched within kinase-specific substrate networks. Null distributions were generated by random sampling of phosphosites with matched study bias, and statistical significance was assessed using the Mann–Whitney U test. Final activity scores were normalized against the null model and used for downstream analyses including correlation with transcriptomic programs.

Structural modeling and protein–DNA docking of CDX2

The 3D structures of CDX2 and its phosphorylated form at serine 176 (CDX2-pS176) were predicted using AlphaFold3. To evaluate the physical interactions between the CD24 enhancer sequence and distinct CDX2 conformations, DNA–protein docking analysis was performed using the HDOCK web server. Docking was conducted separately for unmodified CDX2 and CDX2-pS176 using the enhancer-matched binding motif sequences defined from prior motif analysis.

Inference of copy number variations from single-cell transcriptomes

To infer large-scale copy number variations (CNVs) at single-cell resolution, we applied the inferCNV R package (v1.14.2) to the processed scRNA-seq data. Tumor epithelial cells, derived from scRNA-seq of matched normal colorectal tissue, were analyzed using adjacent normal epithelial cells as reference. Genes with an average expression below 0.1 were filtered out, and denoising steps were applied to reduce technical noise and improve CNV signal clarity.

Meta-cell construction using SEACells

To overcome technical noise and data sparsity inherent in scRNA-seq data, we applied the SEACells algorithm to identify robust transcriptional meta-cell states. A total of 1,700 SEACells were initialized across the dataset, with meta-cell resolution controlled by setting convergence_epsilon = 1e-5, min_iter = 20, and max_iter = 30. Principal component analysis (PCA) was used to construct the cell-cell similarity kernel. The algorithm was run in default mode with adaptive fitting of the SEACells model to stabilize local transcriptomic structure.

Trajectory inference using Palantir and CellRank

Trajectory analysis was performed on single-cell RNA-seq data restricted to cancer cells. Diffusion-based pseudotemporal ordering was computed using the Palantir algorithm. First, diffusion maps were generated using palantir.utils.run_diffusion_maps with 10 components and PCA coordinates derived from UMAP embedding (pca_key = 'X_umap'). Palantir was then run with 500 waypoints and manually defined terminal states using palantir.core.run_palantir. The resulting pseudotime values were used as input for CellRank to compute lineage probabilities. CellRank integrated Palantir pseudotime and transcriptomic similarity into a combined kernel to infer fate probabilities and terminal states. These were used to model cell-state transitions and identify lineage-associated gene programs.

Trajectory-based signature derivation and TCGA deconvolution

Following trajectory inference using Palantir and CellRank, three terminal states (Terminal 1–3) were identified from the pseudotemporal progression of cancer cells. To characterize these terminal states, we constructed gene expression signature matrices for each terminal state using the buildSignatureMatrixMAST function from the DWLS (Digital Weighted Least Squares) R package. These trajectory-derived signatures were subsequently applied to bulk TCGA transcriptomic data using CIBERSORT to infer the enrichment of each terminal program across patient samples.

Pathway enrichment analysis in single-cell and spatial transcriptomics

Pathway-level enrichment analyses at both single-cell and spatial transcriptomic resolution were performed using the escape R package, which facilitates integration of gene expression data with gene set enrichment analysis (GSEA). The tool supports enrichment computation from Seurat and SingleCellExperiment objects and enables access to the Molecular Signature Database (MSigDB v7.0), including Hallmark and other curated gene sets. Enrichment scores were computed for selected gene sets of interest, and pathway activity was assessed across individual cells or spatial regions.

Survival analysis based on omics-derived signatures

Progression-free survival (PFS), and disease-free survival (DFS) analyses were conducted using TCGA colorectal cancer samples. For each molecular signature of interest, patients were stratified into high and low groups based on optimal cutpoints. These cutpoints were computed using the maxstat R package, which identifies the most statistically significant threshold via maximally selected rank statistics. Kaplan–Meier survival curves were generated, and group differences were assessed using the log-rank test.

Construction of response- and PFS-associated gene networks

To investigate functional gene connectivity associated with treatment response and progression-free survival (PFS) at single-cell resolution, we constructed cell-type-specific gene interaction networks using the scHumanNet framework. This method utilizes the SCINET algorithm in conjunction with the HumanNet v3 interactome to infer context-specific gene networks from single-cell RNA-seq data. SCINET incorporates imputation, transformation, and normalization steps via the ACTIONet package to robustly model gene–gene interactions in sparse single-cell data. Subsampling was applied per cluster to stabilize interaction inference across cell types. Response-specific and PFS-specific networks were generated separately. For downstream network characterization, we computed connectivity metrics for marker genes identified in each terminal state, including degree centrality, eigenvector centrality, and PageRank scores. These metrics were used to assess the relative influence of terminal-associated genes within the constructed networks, providing insights into their functional roles in response and survival-associated pathways.

Cell–cell interaction analysis using NICHES in single-cell and spatial transcriptomics

To investigate intercellular communication across tumor and stromal compartments, we used the NICHES framework to compute cell–cell interactions based on known ligand–receptor relationships. In the single-cell RNA-seq dataset, cancer cells were designated as ligand-sending cells, and pairwise interactions with all other cell types were calculated at the individual cell level. For each recipient cell type, interaction scores were derived across all ligand–receptor pairs, and top-ranked interactions were identified using the Seurat

FindMarkers function with a stringent threshold of logfc.threshold = 4.5. In parallel, spatial transcriptomics data were also processed using NICHES, enabling spatially contextualized analysis of cell–cell signaling. For these data, interactions were assessed using spatially co-localized expression patterns, allowing for visualization and inference of neighborhood-level communication.

Ligand–target interaction analysis using NicheNet

To further characterize intercellular communication pathways involving cancer-derived CD24 signaling, we employed NicheNet, a computational framework that integrates ligand–receptor binding, intracellular signaling, and transcriptional regulation into a comprehensive ligand–target prior model. This model provides a quantitative matrix describing the potential regulatory influence of each ligand on downstream target genes based on integrated biological knowledge. As input, cell type–specific marker genes were identified from single-cell RNA-seq data using Seurat's FindMarkers function. We then filtered these marker genes using NicheNet's ligand–target matrix to retain only those with a regulatory score of ≥ 0.0005 for CD24, the ligand of interest. The ligand activity score in NicheNet reflects how well a given ligand is predicted to regulate its downstream target genes, based on prior knowledge from curated signaling and gene regulatory networks. To gain insight into the biological functions of the identified target genes, we performed pathway enrichment analysis using the enrichGO function from the clusterProfiler R package. This allowed us to identify Gene Ontology (GO) terms and biological processes significantly associated with CD24-regulated transcriptional programs in each recipient cell population.

Ligand–receptor scoring in bulk RNA-seq data

Ligand–receptor interaction scores were calculated from bulk RNA-seq data using the decoupleR package. Ligand–receptor pairs were obtained from the LIANA resource and reformatted into gene sets, with each pair treated as a functional unit. Enrichment scores were computed using a weighted mean-based approach. All analyses were conducted using default parameters.

Cell type annotation of spatial transcriptomics data

Cell type deconvolution of the spatial transcriptomics dataset was performed using SPOTlight, which enables reference-based mapping of spatial spots to single-cell identities. A reference atlas was generated from the annotated single-cell RNA-seq dataset of tumor-derived epithelial and stromal cells obtained from the EGAC00001002526 cohort, which had undergone rigorous quality control, normalization, and clustering as described above. SPOTlight was then applied to infer the proportional representation of each cell type within individual spatial transcriptomic spots.

Spatial enrichment analysis of CD24–SIGLEC10 interaction score

To delineate regions with elevated CD24–SIGLEC10 interaction within the spatial transcriptomic landscape, we calculated interaction scores across all spatial spots using the createNumericAnnotations function from the SPATA2 package. Spots with scores exceeding the 80th percentile—derived from the overall distribution of CD24–SIGLEC10 interaction values—were designated as enriched regions. Spatial coherence of these regions was further refined using inner_borders = TRUE and a concavity parameter of 1.5, enabling precise identification of tumor microenvironment niches with potentially high CD24–SIGLEC10 signaling activity.

Spatial niche identification and multi-level signaling landscape analysis

To identify spatially distinct tumor niches and characterize their associated signaling ecosystems, clustering was performed on the spatial transcriptomic data using Seurat. After setting the default assay to SCT (Single-Cell Transform), unsupervised clustering was conducted using FindClusters (resolution = 0.9), followed by UMAP projection (RunUMAP, dims = 1:30) for spatial visualization. For each resulting niche, key signaling axis scores—including those reflecting tumor-intrinsic pathways, intercellular communication, and downstream responses in recipient cell populations—were averaged and quantile normalized to a 0–1 range. To identify niche-specific signaling programs, two-sided t-tests were conducted comparing each niche to all others, with corresponding P values used to determine statistically enriched signaling states.

Projection of single-cell-derived macrophage signatures onto bulk transcriptomes

To investigate the functional distribution of macrophage subtypes at the bulk level, we derived marker gene signatures for resident tissue macrophage-like cells (RTMC) and monocyte-derived macrophages or dendritic cells (moMC/moDC) from single-cell RNA-seq data. Top 50 non-overlapping marker genes were identified for each subtype to ensure subtype-specificity. These gene sets were then used to calculate GSVA enrichment scores for each TCGA-COAD sample, yielding sample-level RTMC and moMC/moDC scores. The samples were projected onto a two-dimensional space defined by these scores, and unsupervised clustering was performed using k-means (k = 4) to stratify the samples into four groups with distinct macrophage-related immunological features.

Fig. 3F

Relative MFI		
Scramble	sgSLC39A4	sgSLC39A5
1	1.076	1.231
1	1.126	1.444
1	0.998	1.146
1	1.073	1.215

Fig. 3G

Relative mRNA level				Relative mRNA level		
	Scramble	sgSLC39A4			Scramble	sgSLC39A5
CD24	1	3.375		CD24	1	4.084
	1	3.110			1	4.029
	1	5.191			1	4.815
	1	2.332			1	4.435
VAV3	1	1.992		VAV3	1	2.486
	1	2.589			1	2.436
	1	1.949			1	1.926
	1	1.616			1	2.379
P21	1	1.436		MUC2	1	2.192
	1	1.546			1	1.925
	1	1.825			1	2.669
	1	1.566			1	2.063
HEPH	1	1.149		P21	1	1.475
	1	1.168			1	1.412
	1	1.130			1	1.876
	1	1.394			1	1.265
MUC2	1	1.595		HEPH	1	1.127
	1	1.087			1	1.066
	1	1.313			1	0.918
	1	0.710			1	1.438
SI	1	1.319		SI	1	1.146
	1	1.210			1	1.579
	1	1.170			1	0.676
	1	0.588			1	0.857
CDH17	1	1.049		CDH17	1	1.163
	1	0.889			1	0.924
	1	1.028			1	1.067
	1	1.148			1	0.940
CA1	1	0.496		CA1	1	0.700
	1	0.249			1	0.498
	1	0.381			1	0.782
	1	0.289			1	0.439

Fig. 3H

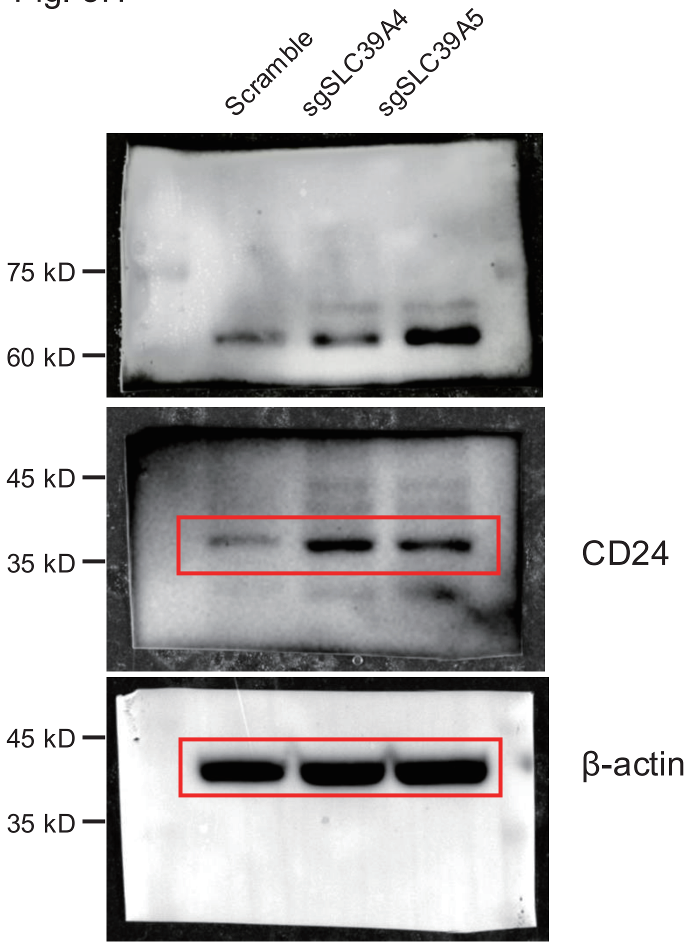

Scramble
sgSLC39A4
sgSLC39A5
75 kD
60 kD
45 kD
35 kD
CD24
45 kD
35 kD
β-actin

Fig. 3I

Relative MFI		
Scramble	sgSLC39A4	sgSLC39A5
1	1.749	1.481
1	2.712	2.107
1	1.647	1.231
1	1.685	1.307

Fig. 3J

	Relative mRNA level		
	Scramble	sgSLC39A4	sgSLC39A5
CD24	1	3.375	4.084
	1	3.110	4.029
	1	5.191	4.815
	1	2.332	4.435
	Relative mRNA level		
	Scramble	sgSLC39A4	sgSLC39A5
CD47	1	1.207	1.129
	1	1.185	1.005
	1	1.131	1.031
	1	1.311	1.242

Fig. 7D

	moMC				RTMC	
	Relative mRNA level				Relative mRNA level	
	moMC	RTMC			moMC	RTMC
CD40	1	0.023		CD206	1	25.688
	1	0.037			1	25.608
	1	0.038			1	7.059
	1	0.021			1	31.004
	1	0.025			1	2.046
	1	0.064			1	3.003
CD64	1	0.109		F13A1	1	565.350
	1	0.103			1	19.041
	1	0.321			1	774.166
	1	0.157			1	520.048
	1	0.125			1	207.145
	1	0.187			1	3.719
CXCL16	1	0.076		SLC40A1	1	4.870
	1	0.109			1	6.180
	1	0.091			1	35.383
	1	0.068			1	5.079
	1	0.077			1	8.022
	1	0.153			1	8.444

Fig. 7E

	Normalized phagocytosis				Normalized phagocytosis	
	Scramble	sgSLC39A4			Scramble	sgSLC39A5
IgG	0.923	0.579		IgG	0.923	0.615
	1.000	0.847			1.000	0.527
	1.000	0.739			1.000	0.637
	1.000	0.705		anti-CD24	0.949	1.000
	1.000	0.582			0.931	0.855
anti-CD24	0.949	0.692			1.000	0.858
	0.931	0.870				
	1.000	0.801				
	0.995	0.921				
	0.865	0.915				

Fig. S3C

Relative mRNA level				Relative mRNA level		
	Scramble	sgSLC39A4			Scramble	sgSLC39A5
SLC39A4	1	30.065		SLC39A5	1	3.931
	1	38.976			1	2.978
	1	24.599			1	2.099
	1	34.190			1	2.373

Fig. S3E

Relative mRNA level		
	Scramble	sgCDX2
CDX2	1	9.536
	1	3.766
	1	6.084

Fig. S3F

Relative mRNA level		
	Scramble	sgCDX2
SI	1	16.842
	1	45.025
	1	20.421
CA1	1	7.890
	1	14.934
	1	18.961
VAV3	1	7.534
	1	6.589
	1	8.953
MUC2	1	5.016
	1	5.712
	1	6.106
CD24	1	4.574
	1	3.524
	1	4.384
HEPH	1	1.892
	1	2.685
	1	3.648
P21	1	1.702
	1	1.322
	1	1.789
CDH17	1	1.569
	1	1.232
	1	1.280