Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Spatial organization of CAF subtypes establishes a VCAN-laden neurogenesis-mimicry barrier driving immune exclusion and immunotherapy resistance in gastric cancer

Laishou Yang
Harbin Medical University Cancer Hospital
Jiale Fan
Harbin Medical University Cancer Hospital
Zhe Liu
Harbin Medical University Cancer Hospital
Junzhu Dai
Harbin Medical University Cancer Hospital
Lei Wang
Harbin Medical University Cancer Hospital
Xiqing Zhu
Harbin Medical University Cancer Hospital
He Yuan
Harbin Medical University Cancer Hospital
Zipu Dong
Harbin Medical University Cancer Hospital
Tonghui Wei
Harbin Medical University Cancer Hospital
Hongjiang Song
Harbin Medical University Cancer Hospital
Rong Huang
Harbin Medical University Cancer Hospital
Linqiang Li
The First Affiliated Hospital of Harbin Medical University
Zhichao Yu
Harbin Medical University Cancer Hospital
Peng Han
Harbin Medical University Cancer Hospital
Tianyi Fang
Harbin Medical University Cancer Hospital

Haibin Song

songhaibin1980@163.com

Harbin Medical University Cancer Hospital

Research Article

Keywords: Gastric Cancer, Neurogenesis Mimicry, Cancer-Associated Fibroblasts, Virtual Knockout/scTenifoldKnk, Risk Score, Immune-excluded microenvironment

Posted Date: October 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11090174/v1

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

Additional Declarations: No competing interests reported.

Abstract

Immune checkpoint blockade only generates durable anti-tumor responses in a minor subset of gastric carcinoma patients; transcriptomic evidence indicates stromal fibroblast remodeling is the core driver of immunotherapeutic failure, yet no defined molecular cascade links extracellular matrix stiffening, neural-like cell reprogramming and poor treatment response. We integrated single-cell profiles, Visium spatial sequencing and bulk transcript datasets covering 757 gastric cancer patient specimens to characterize a fibroblast-restricted neurogenesis mimicry (NM) transcriptional program. Matrix-remodeling $POSTN^+/FAP^+$ and structural $COL1A1^+/COL1A2^+$ fibroblast subsets carried elevated NM module scores; CellChat ligand-receptor mapping proved laminin and CD99 signal cascades restructure intercellular communication within high-NM tissue microregions. Unsupervised clustering built upon the 136 fibroblast-specific NM signature separated all enrolled patients into three subgroups with divergent long-term survival trajectories. Ten independent survival prediction algorithms were benchmarked, and random survival forest (RSF) delivered optimal predictive power with an average C-index of 0.84, with stable performance recorded across eight fully independent validation cohorts. Reverse single-cell projection localized all bulk-derived prognostic risk signals to stromal fibroblast and endothelial cell populations. SHAP interpretable modeling paired with scTenifoldKnk single-gene virtual knockout identified a ten-gene functional module centered on VCAN; these genes exclusively expressed in high-NM fibroblasts and correlated with epithelial-mesenchymal transition, invasive phenotypes and angiogenic activity. After adjusting age, gender and pathological stage covariates, continuous RSF risk scores retained independent prognostic capacity. High-risk lesions recorded elevated TIDE values and M2 macrophage enrichment, consistent with an immune-excluded cold tumor phenotype. Visium spatial mapping localized NM fibroblast hotspots adjacent to disordered microvessels and dense fibrotic stromal borders. Cell-free biophysical binding assays (SPR/BLI) paired with in vivo orthotopic syngeneic models validated the dual oncogenic effects of NM-high fibroblasts: VCAN-rich extracellular matrix forms physical barriers to trap cytotoxic $CD8^+$ T cells, while paracrine VCAN directly binds Integrin β1 to activate AKT/ERK signaling, inducing neuron-like transdifferentiation and accelerating gastric carcinoma progression. Multi-dimensional profiling infers fibroblast-dependent neurogenesis mimicry acts as a druggable stromal driver for gastric cancer progression and immunotherapy resistance, with VCAN-laden matrix walls representing a precise therapeutic target to reverse immune exclusion.

Introduction

Wide-spectrum intratumoral heterogeneity and asymptomatic early progression render gastric cancer a global lethal malignancy, ranking it among the top causes of tumor-associated death worldwide[1]. Molecular subtyping pipelines and immune checkpoint inhibitor regimens have entered routine clinical practice, yet unselected patient groups only achieve low objective response rates that limit long-term therapeutic gains[2]. Tumor cell intrinsic genomic variants cannot fully account for clinical treatment failure; abundant transcriptomic and in vivo data confirm bidirectional signal exchange between malignant epithelium and surrounding stroma jointly mediates therapy resistance[3].

Tumor microenvironment (TME) cell fate reprogramming continuously reshapes tissue cellular composition and spatial architecture, frequently converting immunogenic “hot” tumor lesions into immune-suppressed cold micro-niches. Neurogenesis mimicry (NM) stands out among these reprogramming events: this pathological cascade enables non-neoplastic stromal and tumor cells to reactivate embryonic neural developmental pathways, undergo transdifferentiation, secrete neurotrophic cytokines and drive rigid extracellular matrix (ECM) deposition, all of which accelerate malignant progression[4–9]. Past research simplified intratumoral neural molecular features as passive autonomic nerve infiltration into tumor parenchyma. Updated multi-omics datasets reveal endogenous tissue-wide neural transcriptional programs sustain long-term chemotherapy insensitivity and promote metabolic rewiring plus distant metastatic dissemination[5]. ECM mechanical hardening synchronizes tightly with such cell fate shifts, and matrix-derived mechanical signals jointly modulate cellular stemness and spatial positioning within lesions.

Mechanistically, single-cell resolution mapping cannot pinpoint which TME cell populations initiate NM transcriptional cascades, nor have researchers systematically resolved the biophysical and biochemical signaling axes that reshape tissue-wide intercellular communication downstream of NM activation.

Cancer-associated fibroblasts (CAFs) construct the core structural scaffold of gastric cancer stroma; accumulated work confirms their plastic phenotypes support physical immune barrier formation and block anti-tumor immune cell infiltration[10-12]. Early research treated CAFs as a uniform cell population, a view overturned by modern single-cell RNA sequencing datasets that classify fibroblasts into myofibroblastic, inflammatory and antigen-presenting functional subgroups[13–16]. Conserved developmental pathways, particularly canonical Wnt/β-catenin signaling, control how fibroblasts adopt distinct subtypes. Once activated, CAFs secrete large amounts of matrix proteins, forming dense fibrotic layers that shield tumor cell clusters from cytotoxic T cells[17]. CAF activity is known to couple with ECM remodeling, but whether specific hyperactive fibroblast subclusters trigger NM and assemble impermeable, immunosuppressive matrix barriers has not been established.

To address this, we combined single-cell data from several gastric cancer cohorts, multi-center bulk transcriptomes, and Visium spatial transcriptomics[18,19]. Correlation analysis pointed to hyperactivated Wnt signaling as the upstream trigger specific to this three-CAF triad, adding a subtype-restricted stromal program tied to neural-like reprogramming to current CAF classifications. All NM transcriptional activity concentrates within three Wnt-hyperactivated fibroblast subsets carrying $POSTN^{+}$/FAP^{+} matrix remodeling and $COL1A1^{+}$/$COL1A2^{+}$ structural markers. These three CAF subpopulations restructure tissue-wide intercellular communication via dysregulated laminin and CD99 ligand-receptor pairs[20]. The 136-gene fibroblast NM signature was then used to train random survival forest (RSF) prognostic models, with validation performed in independent cohorts. SHAP analysis attributed the model's predictive signal to VCAN, which emerged as the dominant spatial and functional hub of this oncogenic stromal module[21,22]. CRISPR loss-and-rescue experiments, single-molecule biophysics, and immunocompetent in vivo models together revealed a two-layer pathogenic cascade. In one layer, VCAN-enriched ECM assembled dense physical barriers that held $CD8^{+}$ T cells within the stroma, consistent

with reports that versican restricts T-cell access to the tumor parenchyma[23]. In the other, secreted VCAN ligated Integrin β1, driving AKT/ERK activation and a neuron-like phenotypic shift in gastric epithelial tumor cells. This versican–β1-integrin–neuronal-reprogramming axis has, however, been described only in non-gastric settings[24,25]. Integrated multi-omics data resolve the uncharacterized cellular source of NM in gastric cancer and identify VCAN-rich stromal matrix as a precise intervention target to eliminate stroma-mediated immune exclusion.

Methods

Public Data Acquisition and Pre-processing.

Single-Cell RNA Sequencing Datasets

Two independent stomach adenocarcinoma single-cell RNA sequencing datasets (GSE134520, GSE167297) were downloaded from the TISCH2 database. Only cells isolated from pure tumor tissue were retained for downstream analysis to eliminate non-tissue confounding signals. The integrated dataset contained 56,564 single cells and 19,889 unique genes collected across 15 patients and 20 independent tissue samples.

We applied Seurat v4.3.0 to execute quality filtering, LogNormalize normalization and variance-stabilizing transformation-based screening of the top 2000 highly variable genes. The Harmony algorithm eliminated cross-dataset batch effects using 50 principal components across 30 dimensional spaces. Graph-based clustering (FindNeighbors, FindClusters) and UMAP dimensional reduction generated a complete multicellular tissue landscape. Cell lineage labels followed TISCH2 official annotations; biologically similar subsets were merged, including combining monocytes and macrophages into the Mono/Macro group and aggregating scattered $CD8^+$ subclusters into unified $CD8^+$ T cell populations.

Multivariable Adjustment and Stability Analysis

Multivariate Cox regression models incorporating age, gender and pathological stage as confounding covariates were built using the pooled cohort (n=757) containing 21 core prognostic genes. This statistical design mitigates model overfitting and verifies each gene's independent predictive capacity. Bootstrap LASSO-Cox resampling with 1,000 repeated cycles ran exclusively on the TCGA training dataset. The iteration workflow quantified feature selection frequency and confirmed cross-cohort stability of core signature genes.

Model Calibration and Decision Curve Analysis

Survival probability values for 1-, 3-, and 5-year overall survival were extracted from fitted RSF survival curves. All samples with complete clinical survival endpoints were split into ten equal risk deciles for calibration curve plotting and Brier score calculation against empty null models. Decision Curve Analysis (DCA) quantified net clinical benefit across risk thresholds ranging 0.01–0.50 with a 0.01 step increment.

Two clinical reference strategies (“Treat-All”, “Treat-None”) served as benchmarks to evaluate the RSF model’s practical clinical value.

Cellular Essentiality and Virtual Knockout

DepMap CRISPRi loss-of-function screening data supplied gene dependency metrics to validate core model genes’ biological roles. The scTenifoldKnk pipeline simulated single-gene knockout events restricted to fibroblast cell populations; output manifold alignment distances and dRegulation parameters quantified genome-wide transcriptional shifts triggered by individual gene deletion.

Spatial Transcriptomic Architectures

The spacexr package’s RCTD algorithm completed cell-type deconvolution for each Visium spatial spot, with single-cell transcriptomic profiles set as cell lineage reference anchors[26]. Moran’s I statistic measured spatial autocorrelation across tissue slices. The Getis-Ord Gi* formula mapped transcriptional hot and cold spots across all measured tissue positions:

$$G_i^*(i) = \frac{\sum_j w_{ij}x_j - \bar{x}\sum_j w_{ij}}{s \cdot \sqrt{\frac{n\sum_j w_{ij}^2 - (\sum_j w_{ij})^2}{n-1}}}$$

wij = spatial weight matrix, xj = spot-level gene expression, $\bar{x}$ = global expression mean, s = standard deviation, n = total measured spots. Any spatial unit with Gi* > 1.96 and P < 0.05 was classified as a transcriptional hotspot.

Multiplex Immunofluorescence & Spatial Distance Regression

FFPE gastric cancer tissue slices underwent multiplex immunofluorescence staining to map spatial relationships between stromal subtypes, ECM barriers and cytotoxic immune cell niches linked to immunotherapy response. A multi-antibody panel targeted key cell lineage marker proteins.

ImageJ handled cell segmentation, tissue boundary drawing and cell phenotype identification. Exact Euclidean distances (μm scale) were calculated for every tissue tile to dissect immune exclusion structural patterns. Two independent distance metrics were measured for each cell population:1.Minimal distance between individual VCAN⁺α-SMA⁺ fibroblasts and adjacent CD31⁺ vascular endothelial cells 2.Shortest distance between CD31⁺ endothelial cells and neighboring CD8⁺ T lymphocytes 3.GraphPad Prism consolidated raw distance data and generated all descriptive statistical outputs.All spatial distance values underwent Z-score standardization before univariate logistic regression in R v4.6.1, which tested correlations between standardized spatial metrics and immunotherapy response (responder / non-responder). Regression coefficients, 95% confidence intervals and corresponding P-

values were calculated for each model. Continuous predicted response probabilities were plotted with semi-transparent pink bands marking 95% confidence intervals derived from standard error estimates. All statistical tests adopted two-sided design; statistical significance threshold was set at $P < 0.05$.

Surface Plasmon Resonance (SPR) and Bio-Layer Interferometry (BLI) Assays

We measured direct physical binding kinetics between human recombinant VCAN and Integrin β1 using a Biacore T200 instrument (Cytiva) operated at 25 °C. We covalently coupled recombinant human Integrin β1 onto CM5 sensor chips via standard amine-coupling chemistry. We prepared serial two-fold dilutions of purified recombinant VCAN, ranging from 3.125 nM to 100 nM, in HBS-P running buffer, and injected these samples over sensor surfaces at 30 µL/min flow rate. The association phase lasted 180 s, followed by 600 s of dissociation. We processed reference-subtracted sensorgrams in Biacore Insight Evaluation Software and fitted data with a 1:1 Langmuir binding model to obtain association rate constant (k_a), dissociation rate constant (k_d), and equilibrium dissociation constant (K_D). For independent validation of steady-state binding affinity, we ran parallel bio-layer interferometry (BLI) measurements on an Octet R8 system (Sartorius), using biosensors pre-immobilized with either recombinant Integrin β1 or VCAN protein.

Real-Time Live-Cell Confocal Calcium Imaging

We seeded gastric cancer cells onto 35-mm glass-bottom dishes and treated them with conditioned medium (CM) collected from different CAF groups for 48 h. Cells were loaded with 4 µM Fluo-4 AM (Thermo Fisher Scientific). Loading buffer contained standard physiological salt solution plus 0.02% Pluronic F-127; incubation proceeded for 30 min at 37 °C under dark conditions. We captured dynamic calcium fluorescence transients on a Leica TCS SP8 confocal microscope (excitation 488 nm; emission 505–530 nm), acquiring one frame every second. We triggered cell depolarization by perfusing high-K^+ stimulation buffer (60 mM KCl, 80 mM NaCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM HEPES, 10 mM glucose, pH 7.4). Time-dependent fluorescence intensity (F) was normalized against baseline signal (F_0), and we plotted kinetic traces as ($\Delta F/F_0$).

3D Spheroid Co-Culture, Live-Cell Confocal Tracking, and T-Cell Cytotoxicity

We generated gastric cancer spheroids in ultra-low-attachment round-bottom plates and pre-labeled tumor spheroids using CellTracker Red CMTPX (Invitrogen). We embedded spheroids together with genetically modified CAF populations inside standardized Matrigel in 8-well chamber slides. Primary human $CD8^+$ T lymphocytes were isolated from healthy-donor PBMCs by magnetic negative-selection beads (Miltenyi Biotec). T cells were activated for 72 h using anti-CD3/anti-CD28 Dynabeads (Gibco) supplemented with 100 U/mL recombinant human IL-2, stained with CellTracker Green CMFDA, and seeded into the outer matrix zone.

We captured 3D spatiotemporal migration images every 30 min across a 24–72 h time window on an incubator-integrated Olympus FV3000 confocal microscope. We reconstructed single-cell trajectories

and calculated mean migration velocity, directional persistence, and spheroid penetration depth using Imaris software (Bitplane). We assessed intratumoral cell-killing activity via Caspase-3/7 apoptosis readouts, and characterized T-cell activation phenotypes by flow cytometry detecting CD107a, Granzyme B, Perforin, and IFN-γ.

Orthotopic Syngeneic Gastric Cancer Model and In Vivo Therapy

All animal manipulations complied with institutional policies, and experimental protocols received approval from the Institutional Animal Care and Use Committee of Harbin Medical University. We used 6–8-week-old male immunocompetent 615 mice. After anesthesia, we performed upper midline laparotomy to expose the gastric wall. We prepared cell suspensions mixing murine MFC gastric cancer cells and primary CAFs, either wild-type control or sg-Vcan knockout), suspended in Matrigel, and microinjected the mixture into the gastric subserosal layer.

Mice were randomly assigned to four treatment arms (n=6) for each group):

- Vehicle Control (Isotype control IgG)

- VCAN-Target (sg-Vcan CAFs)

- α-PD-1 monotherapy (anti-mouse PD-1 mAb, clone RMP1-14, Bio X Cell; 200 μg/mouse intraperitoneal injection, twice per week)

- α-PD-1 plus VCAN-Target combination therapy

We monitored tumor size dynamically. On day 21, animals were sacrificed to harvest primary tumors. We measured tumor weight and prepared samples for multiplex immunohistochemistry (CD8, CD86, CD206, Foxp3, VCAN), as well as flow-cytometry profiling of tumor-infiltrating lymphocytes.

Prognostic Feature Selection, RSF Hyperparameter Specification, and Benchmark Statistics

Hierarchical Feature Selection

We started with a 136-gene fibroblast-derived NM signature and performed stepwise filtering. Of these genes, 92 could be reliably detected across bulk transcriptome cohorts. We ran univariate Cox regression on the merged dataset (n=757), yielding 41 risk-associated candidate genes satisfying (P<0.001) and (HR>1). To guarantee cross-dataset robustness, we defined our final 21-gene prognostic panel as the overlapping set of hazard genes reaching significance in four independent cohorts (TCGA, GSE26901, GSE62254, merged cohort; individual-cohort (P<0.0001), merged dataset (P<0.001), (HR}>1).

RSF Model Implementation

We constructed the Random Survival Forest (RSF) model with the $r\ \text{and}\ omF\ \text{or}\ estSRC::rfsrc$ R package. Fixed hyper-parameters were set as: ntree = 1000, nodesize = 5, splitrule = "logrank", seed =

123. We kept hyper-parameters unchanged to fairly benchmark predictive performance alongside 10 other competing survival machine-learning algorithms and to avoid over-optimization toward the training dataset.

Statistical Benchmarking

We compared model predictive performance across training and validation cohorts using the non-parametric Friedman test stratified by dataset (P=0.034). This test supported the overall better predictive performance of the RSF algorithm.

Patient Cohort and Tissue Specimens

All FFPE gastric adenocarcinoma tissue blocks were collected from Harbin Medical University Cancer Hospital. Clinical and pathological tumor stages were graded per AJCC official staging manuals. Any patient receiving preoperative chemotherapy, radiotherapy or targeted therapy was excluded from specimen enrollment. The Institutional Review Board and Ethics Committee of Harbin Medical University Cancer Hospital authorized all experimental protocols. Written informed consent was collected from every patient or their legal guardians prior to tissue acquisition.

Results

Single-cell transcriptomic landscape identifies fibroblasts as the primary neurogenesis mimicry cellular substrate.

Cellular lineage reprogramming within the gastric tumor microenvironment frequently alters cell phenotypic characteristics. This abnormal biological event drives gastric cancer progression and induces therapeutic resistance. To identify the specific cell population responsible for neurogenesis mimicry (NM) at single-cell resolution, we combined two published gastric cancer single-cell RNA sequencing datasets (GSE134520 and GSE167297) covering 56,564 single-cell transcriptomic profiles. We removed cross-experimental batch effects via Harmony algorithm, then performed unsupervised cell clustering and annotated 12 stable, transcriptionally distinct cell subsets. This workflow completely depicted the multicellular cellular landscape of primary gastric tumor tissues (Fig. 1A). UMAP visualization grouped by individual patient samples verified effective removal of technical batch effects (Fig. 1B).

We adopted pie charts to quantify the baseline proportion of each cell lineage, and applied stacked bar graphs to compare cellular component differences across samples and datasets (Fig. 1C, D). We defined canonical marker genes for all annotated cell clusters using signature heatmaps, and conducted cell-type-resolved KEGG pathway enrichment analysis to confirm their distinct biological functions (Fig. 1E, Supplementary Fig. 1A).

A critical unresolved question emerged from this atlas: is NM a unique lineage-restricted transcriptional program, or merely a non-specific mesenchymal response overlapping EMT and wound repair pathways?

We addressed this gap with systematic GSEA across the full single-cell dataset[27]. While NM gene sets shared partial overlap with classic myofibroblast markers, pathway enrichment results proved NM operates as an independent transcriptional module enriched in axon guidance, synapse assembly and neurotrophic factor signaling. When we projected NM gene activity onto all cell populations, ridge plots displayed a heavily right-skewed distribution where fibroblast compartments bore nearly all NM transcriptional signals (Fig. 1G, H). UMAP density contours further concentrated high NM scores exclusively over stromal fibroblast clusters, driven by local upregulation of core neural regulators STAT3, TGFB1 and SEMA4F (Fig. 1I).

To rule out bias from a single statistical tool, we deployed three complementary quantitative approaches to compare cellular composition between high-NM and low-NM microenvironments. Simple bar graphs captured stark gaps in absolute cell numbers and relative proportions between the two stromal states (Fig. 1J). We followed up with MiloR neighborhood abundance testing and Fisher's exact odds ratio bubble plots for orthogonal validation (Fig. 1K, L). All three statistical pipelines converged on one consistent conclusion: fibroblasts act as the central cell population governing NM transcriptional activity within gastric cancer stroma.

CAF subpopulation heterogeneity reveals preferential coupling of neurogenesis mimicry with Wnt signaling and extracellular matrix remodeling.

Recent single-cell sequencing studies have revised the traditional consensus that cancer-associated fibroblasts (CAFs) constitute a homogeneous cell population. Instead, the distinct spatial localization and functional heterogeneity of CAFs govern the formation of tumor extracellular matrix barriers, which further block local anti-tumor immune responses. To clarify whether neurogenesis mimicry (NM) activation is a universal feature of all CAFs or restricted to a tumor-specific fibroblast subpopulation, we performed high-resolution re-clustering on all stromal fibroblast cells. This analysis yielded seven transcriptionally heterogeneous CAF subclusters, designated Fib_Sub1 to Fib_Sub7 (Fig. 2A). The raw cell count of each subcluster showed obvious numerical differences, indicating NM signaling is not uniformly activated across all fibroblast populations (Fig. 2B).

Stacked proportion analysis confirmed three subclusters were significantly enriched in high-NM tumor microregions: $POSTN^+/FAP^+$ matrix-remodeling Fib_Sub1, $COL1A1^+/COL1A2^+$ structural Fib_Sub2, and $RGS5^+/PDGFRB^+$ perivascular Fib_Sub3 (Fig. 2C, Supplementary Fig. 2A). We calculated NM module scores for all CAF subtypes via violin plot analysis; the three enriched subclusters exhibited dramatically higher NM activity than inflammatory and quiescent fibroblast subsets. Marker gene dot plots further verified the unique molecular characteristics of these three CAF subtypes (Fig. 2D, Supplementary Fig. 2B).

We then profiled the seven CAF subclusters against 12 well-recognized fibroblast functional signatures to characterize their biological functions (Fig. 2F). Differential expression heatmap analyses revealed that NM transcriptional activity is tightly associated with aberrant Wnt pathway activation and extracellular matrix remodeling. Notably, stromal regions with high and low NM expression exhibited

distinct functional characteristics (Fig. 2G, H). To further verify canonical Wnt pathway activation in high-NM CAFs through biochemical experiments, we performed subcellular fractionation assays on three independent primary CAF samples derived from patients (CAF1, CAF2, and CAF3). The results showed significant nuclear accumulation of β-catenin in high-NM CAFs compared with their low-NM counterparts (Supplementary Fig. 16A–C). In parallel, Western blot detection confirmed the upregulated expression of multiple matrix remodeling and structural proteins, including VCAN, POSTN, FAP, COL1A1 and COL1A2 in high-NM CAFs (Supplementary Fig. 16D–F). These protein-level findings further validate the transcriptomic results obtained from single-cell sequencing. We adopted subcluster-level Spearman correlation analysis to further validate this regulatory interaction. The NM score of $POSTN^+$ Fib_Sub1 showed the strongest positive correlation with Wnt signaling activity ($P < 0.001$), while this correlation was non-significant in iCAF-like Fib_Sub5 cells (Fig. 2I, J).

Overall, these correlation results indicate abnormal Wnt signaling selectively drives the differentiation of Fib_Sub1/2/3. This regulatory axis further induces neural-like matrix stiffening and large-scale tissue architectural remodeling during gastric tumor progression.

We performed Wilcoxon rank-sum tests and observed higher stemness scores in high-NM CAF populations (Fig. 2K). Waterfall plot analysis of metabolic pathways revealed concurrent upregulation of fatty acid metabolism and oxidative phosphorylation in NM-activated CAFs (Fig. 2L, Supplementary Fig. 2C). The enhanced metabolic activity provides stable energy input, which maintains the stem-cell-like phenotype of these CAFs and sustains continuous extracellular matrix deposition.

Neurogenesis status remodels intercellular communication networks and activates laminin and CD99 signaling pathways.

Reciprocal biochemical and mechanical signal exchange between tumor epithelium and stroma governs the overall progression of gastric malignancy. To clarify how elevated neurogenesis mimicry (NM) reshapes intercellular signal transmission within tumor microenvironments, CellChat software was adopted to systematically compare ligand- receptor interactomes between high-NM and low-NM lesions.Quantification revealed drastically intensified signal exchange frequency and overall interaction strength inside high-NM compartments, as quantified by circular network diagrams and paired bar charts (Fig. 3A, 3B). Heatmaps mapping signal input and output intensity pinpointed CAFs as the central signal hub in NM-active regions (Supplementary Fig. 3A). As shown in Fig. 3C, targeted circle plots focusing on fibroblasts, $CD8^+$ T lymphocytes and monocyte-macrophage populations further reinforced that CAFs dominate bidirectional signal crosstalk in high-NM niches.

All differentially activated signaling pathways were ranked by interaction magnitude across two NM subgroups. Three cascades, LAMININ, CD99 and Midkine (MK), displayed the most selective enrichment within high-NM microenvironments (Fig. 3D). Comparative bubble graphs and bar charts listing the top 20 shifted ligand-receptor pairs captured a consistent molecular shift (Fig. 3E, 3F). High-NM stroma upregulated adhesion-related molecular pairs including LAMB3–CD44, COL4A1–CD44 and FN1–ITGAV, which mediate matrix anchoring and tissue stiffening. In contrast, ligand-receptor axes maintaining

balanced matrix homeostasis and anti-tumor remodeling such as THBS1–CD36 were significantly suppressed.The divergent ligand-receptor interactions between groups translated to measurable gaps in pathway-level communication counts (Fig. 3G). This signal remodeling triggered synchronous transcriptional activation of oncogenic programs in neighboring epithelial and endothelial cells. We profiled 50 canonical Hallmark pathways via heatmap visualization and lollipop differential ranking to identify altered functional programs (Fig. 3H, 3I). EMT, angiogenesis and pro-inflammatory signaling were uniformly hyperactivated in cells residing within high-NM microenvironments.

Altered stromal communication establishes a dual-high inflammatory-fibrotic microenvironment. Mechanical anchoring signals from CAFs harden extracellular matrix, while activated EMT and angiogenic programs jointly sustain tumor cell survival and invasive capacity.

A 136-gene fibroblast-derived neurogenesis signature defines three molecular subtypes with distinct survival outcomes.

Single-cell-level NM features require translation into bulk transcriptomic biomarkers for clinical patient stratification. We built a two-step filtering pipeline to extract fibroblast-specific NM signature genes. Genes retained simultaneously met two thresholds: positive correlation with fibroblast NM scores ($r > 0$) and fibroblast-biased expression ($logFC > 0.2$, adjusted $P < 0.01$). The intersection of two gene lists yielded a total of 136 candidate genes, visualized via Venn overlap analysis (Fig. 4A). From this layered profiling, volcano plots and four-quadrant scatter distributions verified these genes carry strong fibroblast specificity and tight correlation with NM transcriptional activity (Fig. 4B, 4C). A heatmap profiling top 25 signature genes across 12 cell lineages further confirmed minimal expression in non-stromal cells (Fig. 4D; Supplementary Data 1).

We applied the signature to a merged meta-cohort containing 757 gastric cancer bulk transcriptomic samples. Univariate Cox regression plotted on volcano charts demonstrated most signature genes act as risk predictors of unfavorable survival (Fig. 4E; Supplementary Fig. 4A). Forest plots singled out VCAN and FBLN5 as the two most powerful hazard- associated genes (Fig. 4F). Unsupervised consensus clustering based on signature expression split all patients into three non-overlapping subgroups (Cluster A/B/C). Consensus matrix heatmap and PCA scatter plots validated stable subgroup separation (Fig. 4G, 4H).

Quantification revealed Cluster A contained the highest proportion of stage III/IV advanced tumors (Fig. 4I). Corresponding Kaplan–Meier curves illustrated drastically shortened overall survival for Cluster A relative to Cluster C (Fig. 4J; Supplementary Fig. 4B). Multi-dimensional functional profiling via CancerSEA and multiple immune deconvolution algorithms uncovered the intrinsic microenvironmental difference between subgroups (Supplementary Fig. 4C, 4D): high-risk Cluster A presented stroma-abundant, immune-excluded niches filled with activated CAFs, endothelial cells and M2 macrophages, while cytotoxic $CD8^+$ T cells were barred from tumor epithelial nests. To rule out batch interference among different patient cohorts and screen reliable prognostic biomarkers, we conducted univariate Cox regression analysis on four independent gastric cancer datasets. We screened overlapping candidate

genes via Venn diagram analysis and obtained 21 risk-related genes with steady prognostic performance across all cohorts (Fig. 4K). Notably, genes linked to favorable patient prognosis barely overlapped across these cohorts (Fig. 4L). We ultimately adopted this conserved 21-gene signature as the core input feature for our subsequent machine learning prognostic modeling.

The 136-gene fibroblast NM signature captures stromal heterogeneity that dictates patient survival. Its derived three molecular subtypes separate immune-excluded high-risk patients who bear advanced-stage tumors, and the cross-cohort validated 21-gene core enables stable prognostic prediction across independent clinical cohorts.

Random survival forest model construction and external validation across twelve cohorts.

Single-factor regression and stepwise survival models are vulnerable to clinical confounding variables, and overfitting commonly overstates the predictive capacity of gene-based risk signatures. We built a benchmark workflow covering ten mainstream survival machine learning algorithms using the 21 core stromal risk genes, aiming to build a stable risk scoring system free of overfitting bias.As shown in Fig. 5A, a C-index heatmap covering all training and validation cohorts reflected obvious performance gaps among different algorithms. Random Survival Forest (RSF) outperformed all competing models comprehensively. Quantification revealed the RSF model reached an average C-index of 0.84 with the smallest standard deviation, striking a perfect balance between predictive precision and robustness (Fig. 5B). Dumbbell charts measuring generalization error across datasets further substantiated the stable external applicability of RSF; its predictive power only slightly declined when applied to independent cohorts (Fig. 5C; Supplementary Fig. 5A).

We calculated a continuous risk score for every patient based on the finalized RSF model. In TCGA, GSE26901, GSE62254 and the merged training cohort, Kaplan–Meier curves separated patients into distinct survival subgroups with P values below 0.001. Time-dependent ROC curves further validated reliable long-term prognostic discrimination ability (Fig. 5D–O; Supplementary Fig. 5B). Eight fully independent external cohorts were recruited for systematic cross-platform validation, covering both microarray and RNA-seq data from geographically diverse patient populations. Across all validation sets, higher risk scores consistently corresponded to shortened overall survival (Fig. 5P–W; Supplementary Fig. 5C). We mapped the bulk transcriptome-derived risk score back onto single-cell transcriptional profiles to trace the cellular origin of prognostic signals. Violin distribution diagrams concentrated nearly all risk signals within fibroblast and endothelial cell populations (Fig. 5X). This spatial localization proves the model’s prognostic power is fundamentally rooted in malignant stromal remodeling rather than intrinsic tumor epithelial alterations.

The RSF model eliminates overfitting defects of traditional regression methods, delivers stable prognostic stratification across 12 independent cohorts, and its predictive signal specifically mirrors pathological transformation of gastric cancer stroma.

Risk score is an independent prognostic factor associated with an immune-excluded microenvironment.

Multivariate Cox regression was conducted with age, gender and TNM pathological stage as confounding covariates to verify whether the risk score can independently predict patient survival. Pie charts first visualized the basic distribution of clinical indicators among high-risk and low-risk subgroups (Fig. 6A). Building on this finding, multivariate forest plots confirmed the risk score acts as an independent adverse survival predictor (HR = 1.11, 95% CI: 1.07–1.16, P < 0.001) (Fig. 6B; Supplementary Fig. 5A).

Multiple immune deconvolution algorithms were jointly applied to dissect stromal-immune landscapes of patients with divergent risk scores. Rank-biserial correlation heatmaps integrating seven immune quantification tools and chemokine expression profiles jointly indicated high-risk lesions feature a typical immune-excluded cold tumor phenotype (Fig. 6C, 6D). Such tumor tissues accumulate abundant tumor-promoting stromal cells including activated CAFs, pericytes and M2-type macrophages, while cytotoxic $CD8^+$ T cells, NK cells and B lymphocytes are largely excluded from tumor epithelial nests.

TIDE algorithm analysis delivered consistent supporting results. Stacked proportion charts showed immunotherapy non-responders accounted for a much larger proportion in the high-risk group (Fig. 6E). Box plots recorded significantly elevated TIDE_CAF, TIDE_Dysfunction, TIDE_Exclusion and total TIDE scores in high-risk specimens (Fig. 6F–I). Angiogenesis, IPRES and M2 macrophage signatures were simultaneously upregulated, whereas Immunophenoscore (IPS) dropped sharply (Fig. 6J–N). A correlation heatmap further solidified the tight linkage between elevated risk scores and activated inflammatory-stromal programs (Fig. 6O). Correlation analysis uncovered two core molecular defects in high-risk tumors: universal transcriptional suppression of T cell-recruiting chemokines CXCL9, CXCL10 and CXCL13, plus widespread downregulation of HLA I/II antigen presentation machinery (Supplementary Fig. 6A). Meanwhile, high-risk lesions carried lower TMB and MSI values together with weakened stemness signatures, while tumor-driving malignant pathways were broadly upregulated (Supplementary Fig. 7A–K). These molecular defects collaboratively build a dense, impermeable stromal barrier that blocks anti-tumor immune infiltration. The RSF risk score serves as an independent clinical prognostic biomarker. High-risk gastric cancers construct immune-excluded microenvironments via suppressed antigen presentation and impaired T cell recruitment, which explains their poor survival and low responsiveness to immunotherapy; corresponding pharmacogenomic profiles provide targeted intervention clues for such patients (Supplementary Fig. 8, 9).

SHAP analysis nominates a ten-gene stromal functional module and spatial transcriptomics validates discrete fibroblast NM hotspots.

Standard survival machine learning models cannot explicitly display internal weight distribution, which hinders the identification of key stromal genes driving gastric cancer patient risk stratification. To address this analytical limitation, we adopted SHapley Additive exPlanations (SHAP) to quantify gene-level weights in our prognostic risk model.

The SHAP value ranking analysis of all 21 candidate signature genes (Fig. 7A) screened out a core ten-gene stromal regulatory panel, including VCAN, FBLN1, FBLN5, GADD45B, CDH11, NNMT, GLT8D2, SGCD, HTRA1 and COL6A2. This gene panel dominated the model's prognostic prediction performance.

We then used DepMap CRISPR gene dependency datasets to verify whether these ten genes regulate the autonomous proliferation of gastric tumor epithelial cells. Most gastric cancer cell lines presented gene effect density values above −0.5 (Fig. 7B), indicating these genes were not essential for the survival of malignant tumor cells. This result suggested the oncogenic role of the ten genes was independent of cell-intrinsic tumor regulation.

Instead, these genes likely promote tumor progression by remodeling the tumor microenvironmental niche. This hypothesis was supported by cell distribution analysis: dot plot results confirmed all ten genes were specifically expressed in fibroblast populations and NM-high tumor microregions (Fig. 7C, 7D). We further analyzed their biological functions via CancerSEA functional annotation and KEGG metabolic correlation analysis (Fig. 7E, 7F). The heatmap results demonstrated that this ten-gene panel mainly participated in tissue structure remodeling and metabolic reprogramming, with comprehensive regulatory effects on stromal biological phenotypes.

Given the single-cell expression characteristics of this stromal gene panel, we adopted Visium spatial transcriptomics of human gastric tumor tissues to locate their in-situ tissue distribution. Spatial clustering analysis showed tumor fibroblasts formed dense regional aggregation in tumor tissues (Fig. 7G); Getis-Ord Gi* hotspot analysis further marked these fibroblast-enriched spatial domains (Fig. 7H). Cell-cell spatial crosstalk mapping revealed fibroblasts acted as the core signal communication hub in the gastric tumor microenvironment (Fig. 7I). Typical panel biomarkers (VCAN, FBLN1 and NNMT) were completely colocalized with fibroblast-abundant regions (Fig. 7J), which verified the stroma-restricted expression pattern of this gene module at intact tissue level.

We next mapped the spatial distribution of NM phenotype by calculating continuous and quartile-grouped NM scores across all tissue spots (Fig. 7K, 7L). Different from diffuse stromal transcriptional activation, high NM transcriptional signals concentrated in discrete spatial hotspots, which were spatially coupled with disordered microvessels and fibrotic tumor boundary regions. Spatial correlation analysis and cross-cell type overlap verification confirmed this positional enrichment feature (Fig. 7M, 7N). Moreover, lineage colocalization analysis of COL6A2 and FBLN1 across six main tumor tissue cell subsets reproduced this fibroblast-specific distribution pattern (Fig. 7O, 7P) .

Finally, we performed in silico single gene knockout via the scTenifoldKnk workflow to verify whether the ten genes function as an integrated regulatory unit in fibroblasts (Supplementary Fig. 6A–J). Knockout of any single gene within this panel triggered significant disruption of the overall fibroblast stromal transcriptional network. Meanwhile, downstream biological pathways related to vascular contractility, focal adhesion and extracellular matrix assembly were significantly dysregulated after single-gene ablation (Supplementary Fig. 6K–T). Collectively, these computational perturbation experiments proved

this ten-gene panel acts as a spatially colocalized, functionally synergistic module to maintain the oncogenic NM-type stromal tumor niche.

VCAN⁺ matrix walls form a physical barrier to drive immune exclusion and neural transdifferentiation

VCAN⁺ matrix walls form a physical barrier to drive immune exclusion and biochemical neural transdifferentiation. Multi-omics integration and spatial deconvolution analyses pointed to VCAN as the key stromal factor that triggers the neurogenesis-mimicry (NM) transcriptional program. To test whether fibroblast-produced endogenous VCAN is required and sufficient for this phenotypic shift, we first screened VCAN isoform expression in primary High-NM CAFs. The V1 isoform dominated transcript abundance in these cells (Supplementary Fig. 10A). We generated stable CRISPR-Cas9 double-guide knockouts (sgVCAN-1, sgVCAN-2) in three independent patient-derived High-NM CAF primary lines. VCAN mRNA and secreted protein were nearly abolished in knockout cells. When we re-expressed a guide-resistant V1 rescue construct, VCAN secretion returned to levels comparable with unmodified CAFs (Supplementary Fig. 10B–D). We collected conditioned medium (CM) from these edited CAFs for functional assays. Endogenous stromal VCAN was necessary to induce AKT and ERK phosphorylation, as well as upregulate neurotrophic factors proNGF and proBDNF across multiple gastric cancer cell lines, including AGS, MKN-45 and HGC-27 (Supplementary Fig. 10E–G). Loss of VCAN also reduced cancer-cell proliferation and Matrigel invasion capacity, which could be rescued by V1 isoform re-expression (Supplementary Fig. 10H–J).

We next carried out cell-free biophysical binding assays to characterize how VCAN engages its cell-surface receptor. Using surface plasmon resonance (SPR), we measured high-affinity direct binding between purified recombinant VCAN and immobilized Integrin β1 (Fig. 8A). Steady-state SPR saturation fitting supported a high-affinity 1:1 binding mode. Bio-layer interferometry (BLI) gave consistent equilibrium-binding results (Fig. 8C, D). When we treated three gastric cancer cell lines of distinct molecular backgrounds (AGS, MKN-45, HGC-27 with High-NM CAF-derived CM, we observed elevated p-ERK, NGF and BDNF levels. Both neutralizing anti-ITGB1 antibody (α-ITGB1) and CRISPR-Cas9-mediated ITGB1 knockout (sg-ITGB1) blocked downstream ERK phosphorylation and neurotrophin induction (Fig. 8E–G). Short-term time-course stimulation (0–120 min) showed that High-NM CM triggered rapid phosphorylation of FAK, SRC, MEK, ERK and AKT; these early signaling events disappeared almost completely after ITGB1 deletion (Supplementary Fig. 12A–C). In epistasis rescue experiments, expressing constitutively-active AKT (ca-AKT) or MEK1 (ca-MEK1) in sg-ITGB1-deficient cells restored MAP2, NeuN and NSE protein abundance (Fig. 8H–J). These results position the AKT/ERK cascade downstream of Integrin β1 receptor ligation.

To determine whether VCAN-driven neurogenesis mimicry acts as a bona fide coordinated neural-like program that is distinct from classical epithelial-mesenchymal transition (EMT) or generalized stress adaptation, we first assessed neuronal marker expression across three independent patient-derived primary CAF cultures (CAF1, CAF2, and CAF3). As shown by immunoblotting, exposure to High-NM CAF CM induced robust expression of neuronal markers (MAP2, NeuN, and NSE). This effect was abrogated

by dual-guide VCAN knockout (sg-VCAN1, sg-VCAN2), while guide-resistant V1 rescue completely restored the above phenotype (Supplementary Fig. 13A–C). We further conducted global transcriptomic profiling via whole-transcriptome RNA-sequencing on gastric cancer cells treated with CM from Ctrl, High-NM, VCAN-KO, and VCAN-Rescue CAFs. Principal component analysis identified distinct global transcriptional separation among these experimental groups (Supplementary Fig. 13D). High-NM CAF exposure triggered coordinated upregulation of gene modules related to axon guidance, synapse assembly, and neurotransmitter signaling, together with elevated EMT and stemness markers. This multi-faceted transcriptional shift was entirely abrogated by VCAN knockout and fully restored upon VCAN rescue (Supplementary Fig. 13E–M).In functional validation assays, live-cell confocal calcium imaging detected robust, neuron-like intracellular Ca influx transients (ΔF/F) in gastric cancer cells following 60 mM KCl depolarization. Such electrophysiological responses were primed by High-NM CAF CM and strictly relied on stromal VCAN function (Supplementary Fig. 13N).

We then examined whether the VCAN-rich extracellular matrix creates physical barriers that suppress cytotoxic-lymphocyte function. We set up a 3D live-cell confocal co-culture system: gastric tumor spheroids were embedded within Matrigel together with different CAF subsets. Over 24–72 hours, High-NM CAFs strongly impaired the motility, migration speed, matrix penetration depth and spheroid infiltration of activated human $CD8^+$ T cells. T lymphocytes became trapped in peripheral matrix regions, leading to reduced tumor-spheroid apoptosis (Supplementary Fig. 14A–I). After VCAN deletion in CAFs, this dense matrix meshwork collapsed. T cells regained directional migration toward tumor spheroids, formed productive contacts with tumor cells, and released cytotoxic effector molecules including CD107a, Granzyme B, IFN-γ and Perforin (Supplementary Fig. 14J). Multiplex immunofluorescence staining of clinical gastric-cancer specimens further supported our model. Dense $VCAN^+$α-SMA^+ matrix sheaths wrapped around microvessels, physically segregating $CD8^+$ T cells away from tumor nests, and this matrix signature correlated negatively with clinical immunotherapy responses (Supplementary Fig. 17A–C).

Lastly, we tested therapeutic strategies targeting this stromal barrier in immunocompetent mice using an orthotopic syngeneic gastric-cancer model. We co-implanted murine MFC gastric-cancer cells together with primary CAFs (Fig. 8K). Due to abundant VCAN-positive matrix deposition, anti-PD-1 monotherapy produced only modest anti-tumor effects. Combining anti-PD-1 treatment with VCAN-deficient sg-Vcan CAFs yielded substantial tumor shrinkage and significantly lower primary-tumor weight ($P < 0.001$) (Fig. 8K–M). Immunohistochemistry confirmed that VCAN ablation plus α-PD-1 treatment dismantled peritumoral VCAN matrix barriers. This combination increased intratumoral $CD8^+$ T-cell infiltration, shifted macrophage polarization toward the M1 phenotype (higher CD86, lower CD206), and reduced $Foxp3^+$ regulatory T-cell populations (Fig. 8O). Tumor-infiltrating-lymphocyte flow cytometry detected more total $CD8^+$ T cells (Supplementary Fig. 15A, B), alongside a marked rise of polyfunctional IFN-γ^+Granzyme B^+ double-positive cytotoxic T cells in the combination-treatment group (Supplementary Fig. 15C, D). Collectively, these observations indicate that matrix-derived VCAN plays dual roles: it mediates physical immune exclusion while triggering biochemical neurogenesis-mimicry programs. Disrupting VCAN-rich matrix walls can relieve immunotherapy resistance in gastric cancer.

Discussion

Although our electrophysiological calcium imaging and RNA-seq demonstrate functional excitability and neural-like gene induction, this transdifferentiation represents a coordinated activation of neural developmental programs (or neurogenesis mimicry plasticity) rather than terminally differentiated mature neurons. Neurogenesis mimicry (NM) remodels gastric cancer tissue spatial architecture and drives both disease progression and poor immunotherapy response. Single-cell sequencing, Visium spatial profiling, machine learning modeling and cellular/in vivo functional assays together delineate the full stromal regulatory cascade uncovered in this work. NM transcriptional activation is not a universal mesenchymal response shared by all fibroblasts; Wnt-hyperactivated $POSTN^+/FAP^+$ matrix remodeling and $COL1A1^+$ structural fibroblast triads exclusively execute this transcriptional program. Three fibroblast subpopulations generate extracellular matrix abundant in VCAN, which drives two separate oncogenic outcomes. The matrix meshwork builds physical barriers that block infiltration of cytotoxic immune lymphocytes. Meanwhile soluble VCAN released by these fibroblasts transmits paracrine signals, switching gastric epithelial cells toward neuron-like phenotypes.

Scholars have long disputed which cell lineage generates neural molecular markers inside gastrointestinal and other solid tumors[28-31]. Previous studies framed intratumoral neural molecular features as a byproduct of peripheral autonomic nerve infiltration[32–34]. Drawing on integrated multi-omic single-cell and spatial datasets, our work revises this narrow view: resident stromal fibroblasts, rather than invading neural tissue, stand as the central cell population that drives the full neurogenesis mimicry (NM) transcriptional cascade.

Numerous single-cell profiling studies have demonstrated that cancer-associated fibroblasts (CAFs) carry divergent abilities to remodel extracellular matrix (ECM) and inhibit anti-tumor immune activity[11-13]. Independent publications have separately characterized $POSTN^+$ and $COL1A1^+$ fibroblast populations across solid malignancies, yet no prior work has clarified if these two subsets coordinate NM activation as a unified functional group. Laminin and CD99 have been described as matrix adhesion molecules across several tumor types[20]. In our CellChat analysis, both fall within a single signaling module selectively active in NM-enriched stroma. High-NM regions show increased LAMB3–CD44 and FN1–ITGAV pairing, which strengthens cell-matrix adhesion, raises ECM stiffness, and confines epithelial clusters within fibrotic pockets.

Strong Wnt signaling drives $POSTN^+/FAP^+$/COL1A1 CAFs toward this specialized fate. NM-active fibroblasts also upregulate oxidative phosphorylation and peroxisomal metabolism; this metabolic shift supports their stem-like state and continuous matrix production, a vulnerability not seen in quiescent or inflammatory fibroblasts.

Most published gastric cancer prognostic signatures are derived from epithelial gene expression. By contrast, our random survival forest (RSF) model, built on fibroblast NM-related genes, performs consistently well across twelve independent patient cohorts [35,36], offering a stroma-centered alternative for risk stratification. Compared with classical Cox proportional hazards regression, which

requires the proportional hazards assumption, the RSF model is non-parametric and can model non-linear, time-dependent hazard patterns across different clinical datasets.

Two interconnected regulatory axes govern subtype-restricted NM activity: upstream Wnt-mediated transcriptional reprogramming, and downstream rewiring of cellular metabolism. Within NM-positive regions, the rewired signaling forms a self-reinforcing oncogenic loop. Laminin and CD99 adhesion signals are amplified, whereas homeostatic pairs such as THBS1–CD36 are suppressed. The imbalance pushes epithelial and endothelial cells into a co-activated inflammatory-fibrotic state, with EMT and angiogenesis programs running in parallel.

3D live-cell tracking shows that dense VCAN-rich matrix directly slows $CD8^+$ T cell movement. This physical barrier works together with reduced chemokine recruitment (lower CXCL9/CXCL10) and M2 polarization to maintain the immune-excluded state. Combining SHAP with Visium mapped VCAN as the central effector of the ten-gene stromal module, with NM fibroblast hotspots sitting next to disorganized microvessels and fibrotic tumor margins. Biophysical binding, functional perturbation, and multiplex immunofluorescence support a two-layer model: VCAN-rich ECM ensheaths $CD31^+$ microvessels and retains $CD8^+$ T cells in peripheral stroma, while soluble VCAN engages Integrin β1 on tumor cells to activate AKT/ERK, raise NGF and BDNF, and push cells toward a neuron-like phenotype.

Projecting fibroblast-specific single-cell signatures onto bulk transcriptomes offers a practical route from high-resolution atlases to clinical risk scores. The 136-gene NM signature separates gastric cancer patients into three groups with different survival and immune microenvironment profiles, complementing existing epithelium-based molecular classifications. After adjustment for age, sex and TNM stage, the RSF risk score independently predicts poor outcome. High-risk tumors downregulate T cell-recruiting chemokines (CXCL9, CXCL10, CXCL13) and HLA antigen presentation, consistent with the low checkpoint-blockade response in unselected gastric cancer. Targeting VCAN or its downstream Integrin β1–AKT/ERK axis could relieve this stromal barrier and restore checkpoint sensitivity, a stroma-directed strategy for chemo- and immunotherapy-refractory disease[37].

Several caveats apply. All Visium data come from one gastric tumor, so we cannot tell whether NM hotspot patterns differ between intestinal and diffuse subtypes. The clinical cohort excludes patients who had neoadjuvant chemotherapy or radiotherapy, so how such treatment remodels NM fibroblasts and VCAN matrix remains unknown. Public DepMap CRISPR screening resources only include epithelial tumor cell lines, leaving the intrinsic survival requirements of the ten core stromal genes uncharacterized in primary human fibroblasts.

Future research directions fall into four clear branches. First, prospective clinical cohorts will validate the predictive value of our RSF risk score to guide immunotherapy regimens. Second, humanized immune mouse models can test combined therapy using anti-VCAN neutralizing antibodies and metabolic inhibitors specific to NM-high CAFs, measuring combined anti-tumor and immune-releasing effects. Third, multi-omics spatial profiling integrating proteomics and chromatin accessibility data will uncover epigenetic and post-translational mechanisms controlling VCAN secretion in $POSTN^+$/FAP^+ fibroblasts.

Fourth, large retrospective clinical cohorts with mIF staining can calculate statistical correlations between stromal VCAN density, NM hotspot abundance and long-term overall survival plus immunotherapy response endpoints.

Integrated multi-omics, spatial mapping and functional experimental data infer Wnt-activated $POSTN^{+}/FAP^{+}$ fibroblast neurogenesis mimicry acts as a core stromal driver of immune exclusion and gastric cancer progression, with VCAN-laden extracellular matrix walls representing a tractable microenvironmental therapeutic target to overcome multi-modal treatment resistance[38-40].

Declarations

Funding declaration

This work was supported by the National Natural Science Foundation of China (Grant No. 82403568) to Tianyi Fang, by grants (PL2025H169 and PL2025H189) from the Joint Fund Cultivation Project of Heilongjiang Provincial Natural Science Foundation to Haibin Song and Peng Han, and by the Harbin Medical University Cancer Hospital Ascend Leading Disciplines Plan (Grant No. PDYS-2024-14) to Peng Han.

Acknowledgements

We gratefully acknowledge all providers of public data used in this study for their contributions.

Author contributions

LY, TF and HS conceived and designed the study. LY, TF, and HS supervised the project and provided administrative support. LY, JF, ZL, JD, XZ, LW, and PH performed the wet-lab functional validation and clinical multiplex immunofluorescence staining experiments. LY, FT, ZL, HY, ZD, TW, and RH developed the bioinformatics pipelines and executed the multi-omics data integration, machine learning modeling, and spatial transcriptomics analysis. LY, JF, XZ, and HS provided clinical specimens and curated the patient repository data. LY, JF, ZD, and RH generated the figures and performed statistical modeling. LY drafted the original manuscript. TF and HS critically revised and edited the manuscript for intellectual content. All authors read and approved the final manuscript.

References

1. Sung H, Filho AM, Laversanne M, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. CA Cancer J Clin. 2026;76(4):e70090.
2. Janjigian YY, Shitara K, Moehler M, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal

adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. Lancet. 2021;398(10294):27-40. doi:10.1016/S0140-6736(21)00797-2

3. Sahai E, Astsaturov I, Cukierman E, et al. A framework for advancing our understanding of cancer-associated fibroblasts. Nat Rev Cancer. 2020;20(3):174-186.
4. Kumar V, Ramnarayanan K, Sundar R, et al. Single-Cell Atlas of Lineage States, Tumor Microenvironment, and Subtype-Specific Expression Programs in Gastric Cancer. Cancer Discov. 2022;12(3):670-691.
5. Amit M, Takahashi H, Dragomir MP, et al. Loss of p53 drives neuron reprogramming in head and neck cancer. Nature, 2020; 578(7795): 449–454.
6. Silverman, DA, Martinez, VK, Dougherty, PM, et al. Cancer-Associated Neurogenesis and Nerve-Cancer Cross-talk. Cancer Res. 2021; 81 (6): 1431-1440.
7. Bloomer H, Dame HB, Parker SR, Oudin MJ. Neuronal mimicry in tumors: lessons from neuroscience to tackle cancer. Cancer Metastasis Rev. 2025;44(1):31.
8. Lu R, Fan C, Shangguan W, et al. Neurons generated from carcinoma stem cells support cancer progression. Signal Transduct Target Ther. 2017;2:16036.
9. Sattler AL, Diba P, Hawthorne K, et al. Sympathetic nerve-fibroblast crosstalk drives nerve injury, fibroblast activation, and matrix remodeling in pancreatic cancer. JCI Insight. 2026;11(7):e192814.
10. Aysit N, Altintas E, Koksalar Alkan F, Ozturk G, Korkaya H. The neural niche in cancer: mechanistic insights into tumor-neuron-immune crosstalk and therapeutic opportunities. Front Cell Dev Biol. 2026;14:1667459.
11. Galbo PM, Zang X, Zheng D. Molecular features of cancer-associated fibroblast subtypes and their implication on cancer pathogenesis, prognosis, and immunotherapy resistance. Clinical Cancer Research, 2021; 27(9): 2636–2647.
12. SOzmen E, Demir TD, Ozcan G. Cancer-associated fibroblasts: protagonists of the tumor microenvironment in gastric cancer. Front Mol Biosci. 2024;11:1340124.
13. Öhlund D, Handly-Santana A, Biffi G, et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. J Exp Med. 2017;214(3):579-596.
14. Elyada E, Bolisetty M, Laise P, et al. Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. Cancer Discov. 2019;9(8):1102-1123.
15. Zhang X, Ren B, Liu B, et al. Single-cell RNA sequencing and spatial transcriptomics reveal the heterogeneity and intercellular communication of cancer-associated fibroblasts in gastric cancer. J Transl Med. 2025;23(1):344.
16. Chen J, Chen R, Huang J. A pan-cancer single-cell transcriptional analysis of antigen-presenting cancer-associated fibroblasts in the tumor microenvironment. Front Immunol. 2024;15:1372432.
17. Yang Y, Sun H, Yu H, et al. Tumor-associated-fibrosis and active collagen-CD44 axis characterize a poor-prognosis subtype of gastric cancer and contribute to tumor immunosuppression. J Transl

Med. 2025;23(1):123.
18. Ståhl PL, Salmén F, Vickovic S, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science. 2016;353(6294):78-82.
19. Cheng K, Xu J, Yang J, Zhang Z. Spatial atlas of gastric cancer reveals macrophage-driven stromal activation via ICAM1⁺ smooth muscle cells. BMC Cancer. 2026;26(1):466.
20. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021;12(1):1088.
21. Lundberg SM, Erion G, Chen H, et al. From Local Explanations to Global Understanding with Explainable AI for Trees. Nat Mach Intell. 2020;2(1):56-67.
22. Hu B, Han Z, Niu Y, Wang Y, Shen B, Zhu J. Bulk Sequencing Combined With Single-Cell Sequencing Identifies High Expression of VCAN in Fibroblasts Promoting the Progression of High-Stemness Gastric Adenocarcinoma Cells. FASEB J. 2026;40(2):e71349.
23. Hirani P, McDermott J, Rajeeve V, et al. Versican Associates with Tumor Immune Phenotype and Limits T-cell Trafficking via Chondroitin Sulfate. Cancer Res Commun. 2024;4(4):970-985.
24. Wu Y, Chen L, Zheng PS, Yang BB. beta 1-Integrin-mediated glioma cell adhesion and free radical-induced apoptosis are regulated by binding to a C-terminal domain of PG-M/versican. J Biol Chem. 2002;277(14):12294-12301.
25. Wu Y, Sheng W, Chen L, et al. Versican V1 isoform induces neuronal differentiation and promotes neurite outgrowth. Mol Biol Cell. 2004;15(5):2093-2104.
26. Cable DM, Murray E, Zou LS, et al. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol. 2022;40(4):517-526.
27. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102
28. Tang PC, Chung JY, Liao J, et al. Single-cell RNA sequencing uncovers a neuron-like macrophage subset associated with cancer pain. Sci Adv. 2022;8(40):eabn5535.
29. Mauffrey P, Tchitchek N, Barroca V, et al. Progenitors from the central nervous system drive neurogenesis in cancer. Nature. 2019;569(7758):672-678.
30. Xue M, Zhu Y, Jiang Y, et al. Schwann cells regulate tumor cells and cancer-associated fibroblasts in the pancreatic ductal adenocarcinoma microenvironment. Nat Commun. 2023;14(1):4600.
31. Wang W, Li L, Chen N, et al. Nerves in the Tumor Microenvironment: Origin and Effects. Front Cell Dev Biol. 2020;8:601738.
32. Magnon C, Hall SJ, Lin J, et al. Autonomic nerve development contributes to prostate cancer progression. Science. 2013;341(6142):1236361. doi:10.1126/science.1236361
33. Wang L, Xu J, Xia Y, et al. Muscarinic acetylcholine receptor 3 mediates vagus nerve-induced gastric cancer. Oncogenesis. 2018;7(11):88.

34. Kobayashi H, Iida T, Ochiai Y, et al. Neuro-Mesenchymal Interaction Mediated by a β2-Adrenergic Nerve Growth Factor Feedforward Loop Promotes Colorectal Cancer Progression. Cancer Discov. 2025;15(1):202-226.
35. Xiang YH, Mou H, Qu B, Sun HR. Machine learning-based radiomics score improves prognostic prediction accuracy of stage II/III gastric cancer: A multi-cohort study. World J Gastrointest Surg. 2024;16(2):345-356.
36. Li F, Feng Q, Tao R. Machine learning-based cell death signature for predicting the prognosis and immunotherapy benefit in stomach adenocarcinoma. Medicine (Baltimore). 2024;103(10):e37314.
37. Addala V, Newell F, Pearson JV, et al. Computational immunogenomic approaches to predict response to cancer immunotherapies. Nat Rev Clin Oncol. 2024;21(1):28-46.
38. Liu X, Yao Q, Wang Z, et al. POSTN+ CAFs facilitate gastric cancer peritoneal metastasis by promoting ICAM-1-dependent tumor cell adhesion and CD8+ T-cell exhaustion. Front Immunol. 2026;17:1796080.
39. Hirani P, Gauthier V, Allen CE, Wight TN, Pearce OMT. Targeting Versican as a Potential Immunotherapeutic Strategy in the Treatment of Cancer. Front Oncol. 2021;11:712807.
40. Li D, Tong G, Huang W. Single cell analysis reveals cancer associated fibroblast subtypes as prognostic markers and therapeutic targets in gastric cancer. Discov Oncol. 2025;17(1):182.

Figures

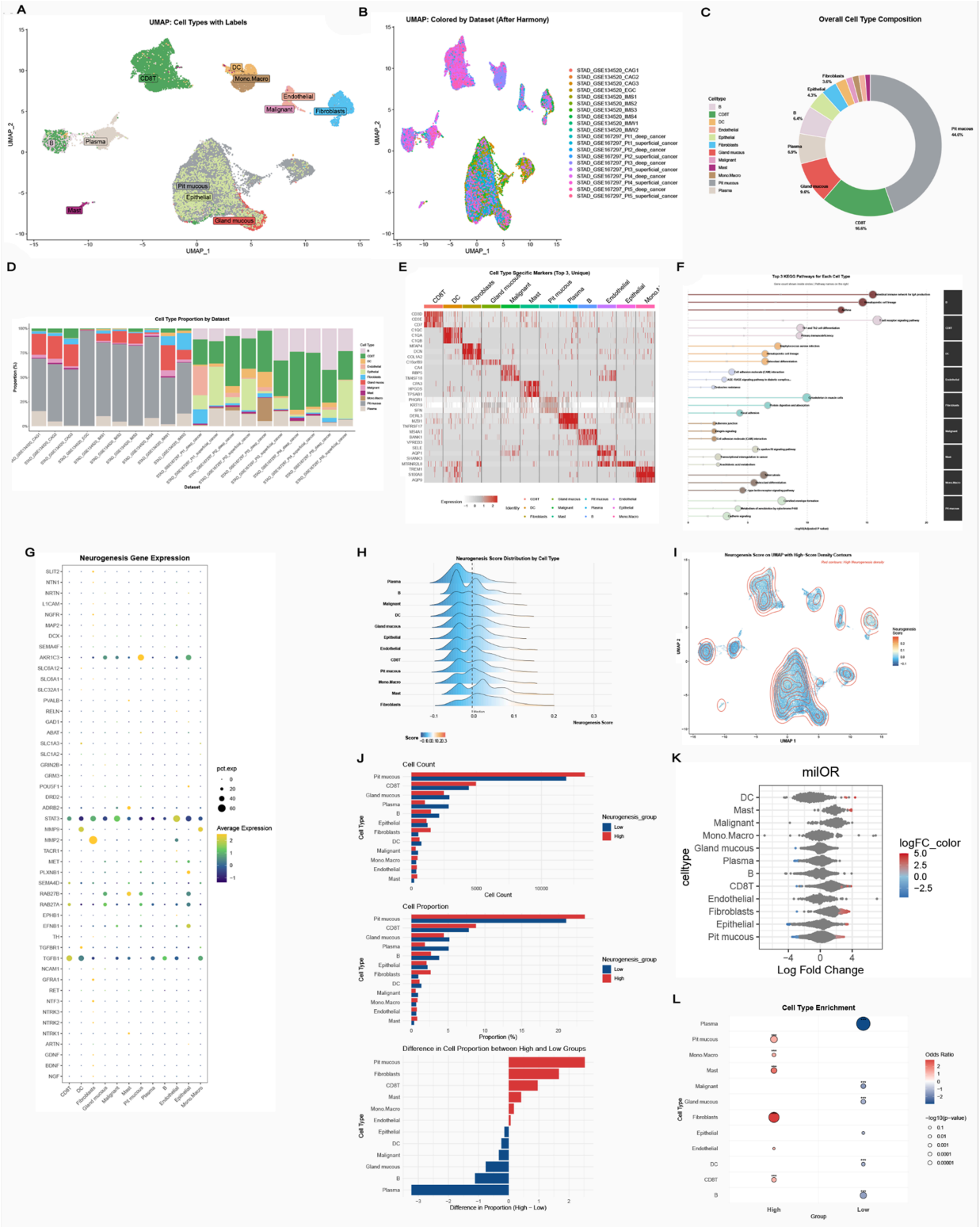


Figure 1

Single-cell transcriptomic landscape of gastric cancer and identification of fibroblasts as the primary neurogenesis mimicry (NM) cellular substrate.

(A) UMAP projection marking 12 annotated cell lineages from integrated single-cell data (B) UMAP layout shaded by sample source, verifying effective batch removal after Harmony integration (C) Pie

chart showing the overall proportional makeup of all detected cell populations (D) Stacked bar graphs tracking cell subtype ratios across individual samples and datasets (E) Heatmap displaying normalized expression of top three marker genes for each cell lineage (F) Dot plot summarizing the three most enriched KEGG pathways per cell group (G) Dot plot comparing global neurogenesis signature expression across distinct cell types (H) Ridge plot quantifying NM module scores for every annotated lineage (I) UMAP density contour map highlighting spatial accumulation of high NM transcriptional signals (J) Bar charts illustrating differences in absolute cell counts and proportional fractions between high- and low-NM microenvironments (K) MiloR violin plot showing neighborhood-level differential cellular abundance (L) Bubble plot presenting odds ratios for cell subset enrichment in high-NM tumor compartments

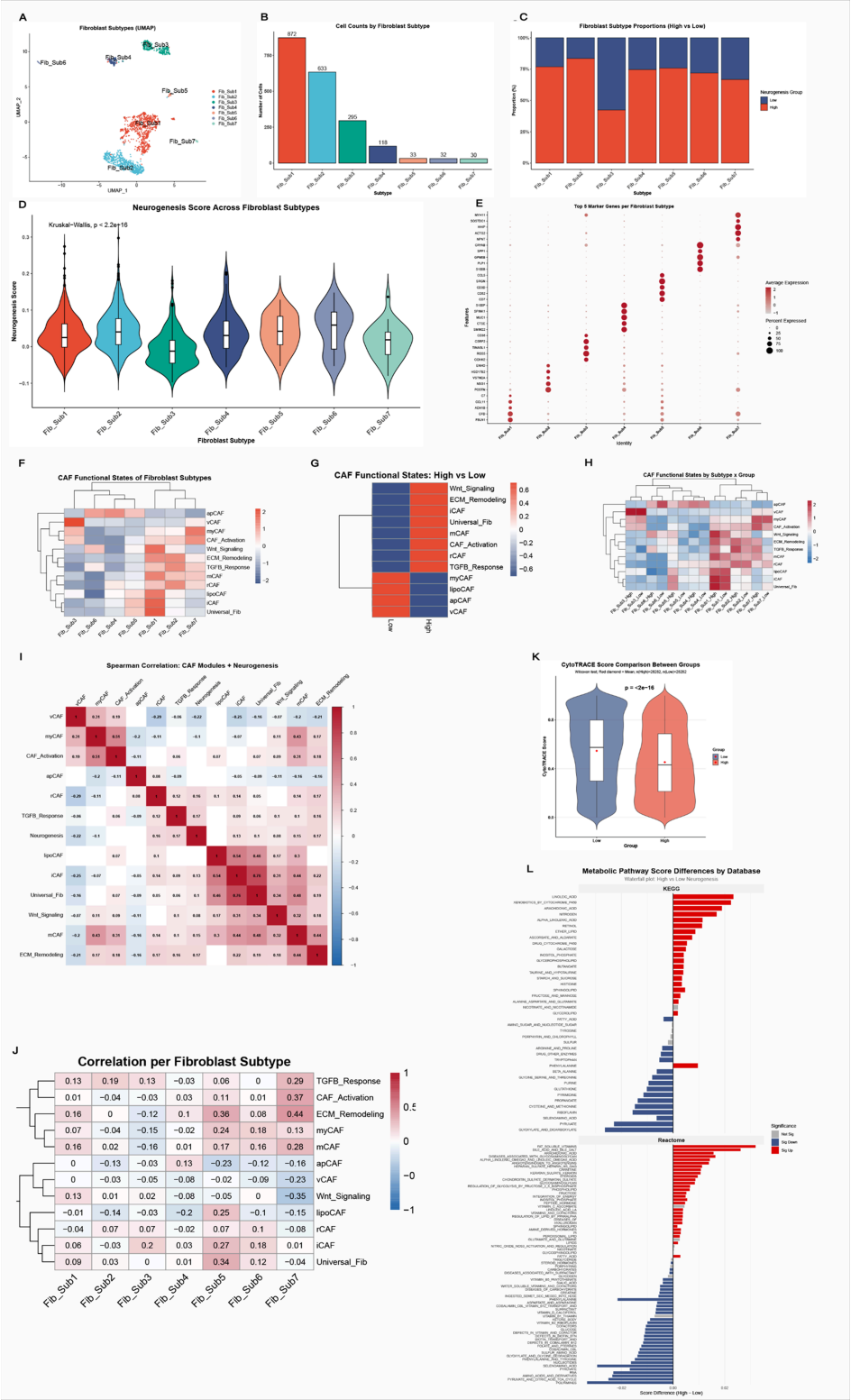


Figure 2

CAF subpopulation heterogeneity and its functional coupling with neurogenesis mimicry.

(A) UMAP plot showing seven transcriptionally distinct fibroblast subclusters (Fib_Sub1–7) (B) Bar graph summarizing total cell abundance of each CAF subgroup (C) Stacked bar diagram comparing CAF subtype proportions between high-NM and low-NM tumor microenvironments (D) Violin plots illustrating

NM module score distribution across seven fibroblast subpopulations (E) Dot plot displaying the top five canonical marker genes for each CAF subtype (F) Heatmap displaying activity scores of 12 canonical CAF functional modules across subclusters (G) Heatmap showing differential CAF module activity between high- and low-NM stroma (H) Combined heatmap integrating CAF functional scores grouped by fibroblast subtype and NM status (I) Spearman correlation heatmap linking CAF functional modules to NM scores (J) Intra-subcluster correlation heatmaps mapping relationships between CAF programs and NM activity (K) Wilcoxon test-based comparison of stemness scores between high- and low-NM fibroblast groups (L) Waterfall plot highlighting metabolic pathways significantly upregulated in high-NM fibroblast populations

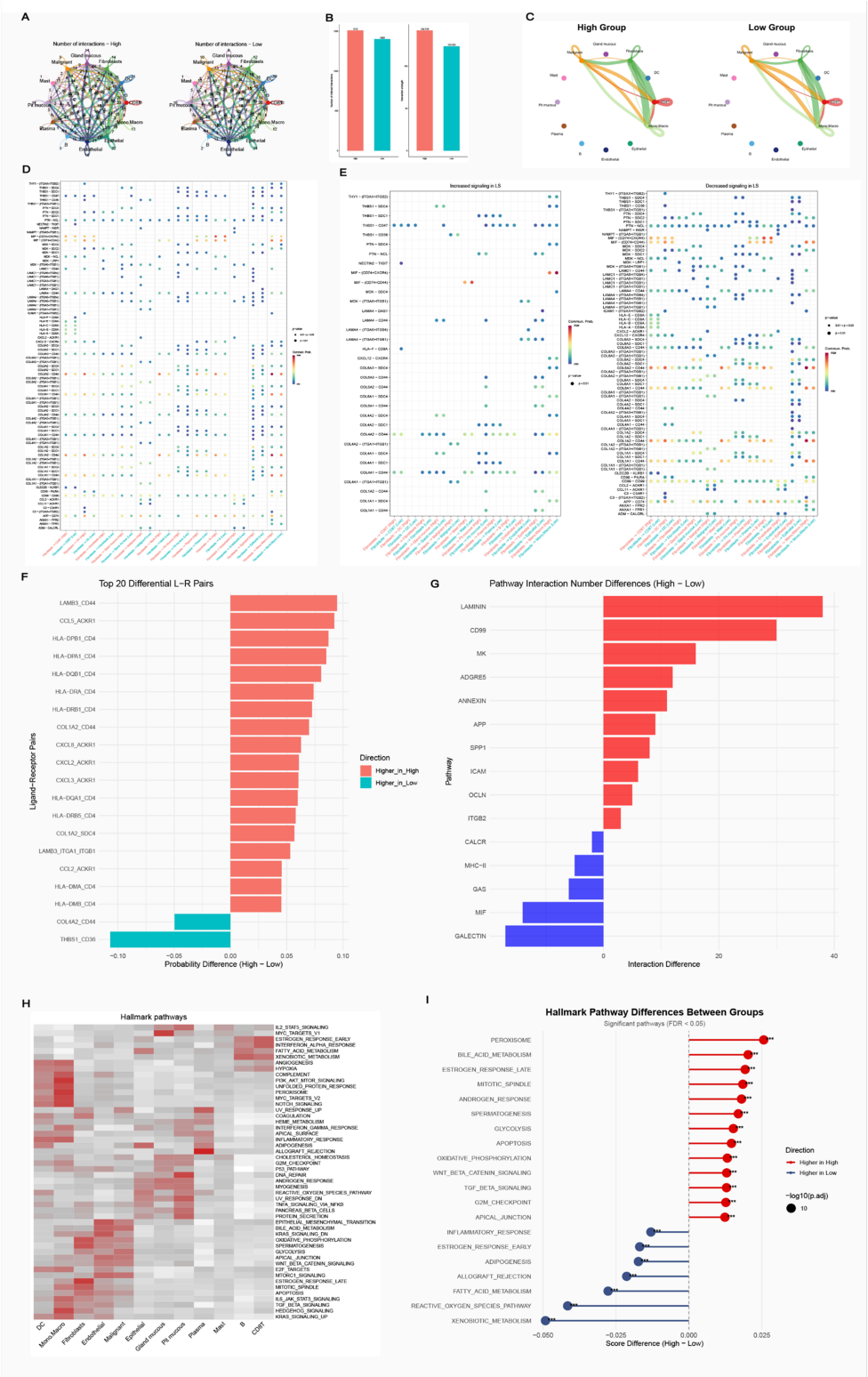


Figure 3

Neurogenesis status remodels intercellular communication networks and activates distinct signaling pathways.

(A) Circular network graphs comparing total ligand-receptor interaction counts between high- and low-NM microenvironments (B) Paired bar charts quantifying total interaction number and signal

intensity across two NM subgroups (C) Focused circular plots displaying signal strength among fibroblasts, $CD8^+$ T cells and monocyte/macrophage populations (D) Ranking of differentially enriched signaling axes identified through CellChat ligand-receptor analysis (E) Bubble graphs separating upregulated and downregulated ligand-receptor pairs in high-NM stroma (F) Bar chart listing the top 20 ligand-receptor pairs with the most significant expression changes (G) Bar graph counting total pathway-level communication events in high versus low neurogenesis mimicry compartments (H) Heatmap displaying activity scores of 50 Hallmark functional pathways across all cell subtypes (I) Lollipop diagram highlighting Hallmark pathways with the largest activation differences between high- and low-NM groups

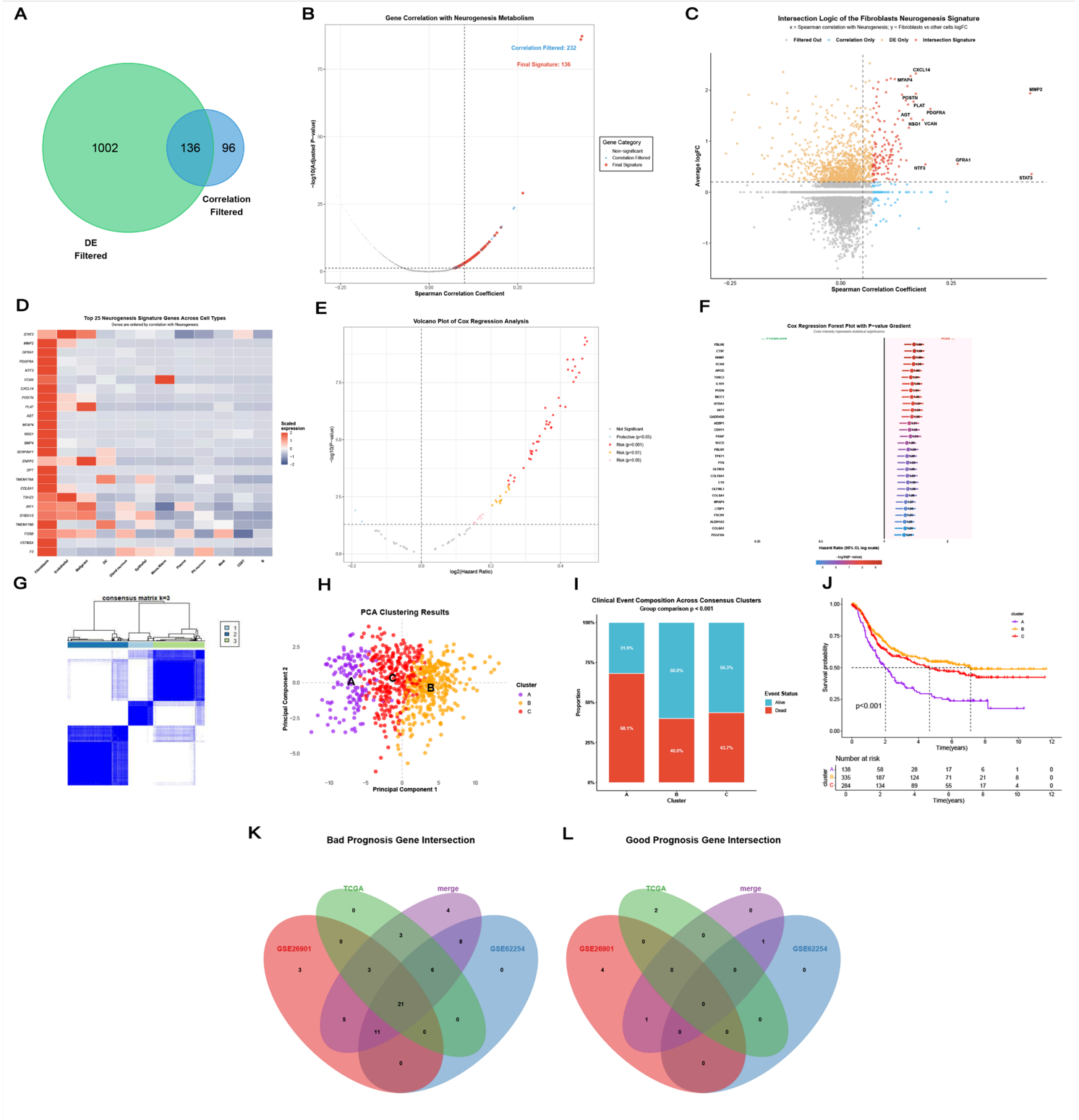


Figure 4

Fibroblast-derived neurogenesis gene signature, consensus molecular classification, and multi-cohort prognostic risk gene intersection.

(A) Venn graph showing overlapping genes from correlation screening and differential expression filtering to form the 136-gene NM signature (B) Volcano plot mapping Spearman correlation coefficients

and statistical significance for all screened genes (C) Scatter plot integrating gene-fibroblast NM correlation and average expression logFC; upper-right quadrant genes compose the final signature (D) Heatmap showing normalized expression of top 25 signature genes across 12 distinct cell lineages (E) Volcano plot summarizing univariate Cox regression outcomes of signature genes in bulk gastric cancer datasets (F) Forest plot highlighting signature genes with the strongest prognostic hazard effects (G) Consensus matrix heatmap supporting stable classification into three molecular subtypes (k = 3) (H) PCA scatter plot colored by the three consensus molecular subtypes (I) Stacked bar chart displaying the distribution of clinical tumor stages across three subtypes (J) Kaplan–Meier overall survival curves comparing clinical outcomes of Cluster A, B and C (K) Venn diagram illustrating 21 consistently enriched risk genes shared across four independent patient cohorts (L) Venn graph showing minimal overlap of protective signature genes among the four validation cohorts

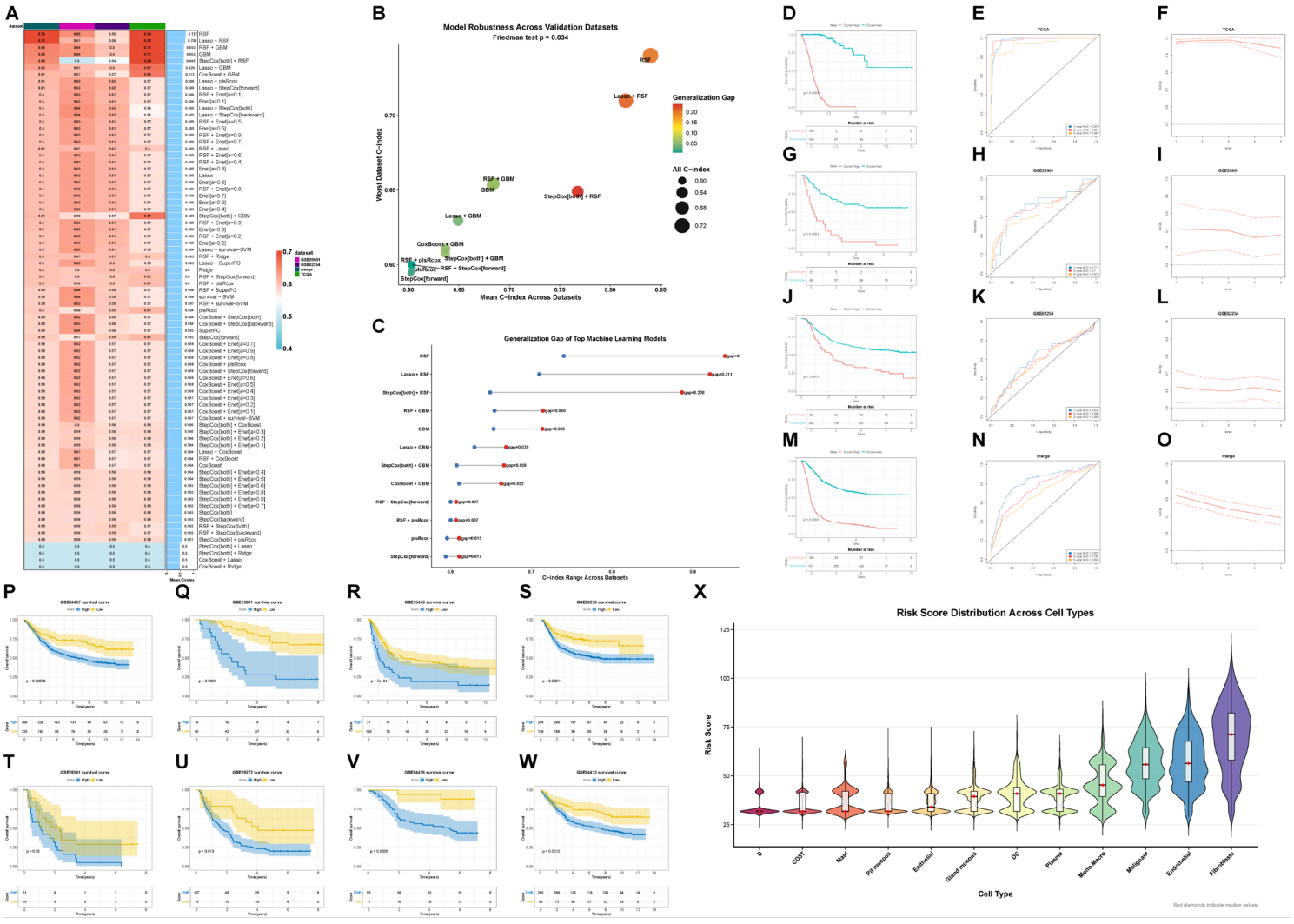


Figure 5

Machine learning model construction, comparison, and external validation.

(A) Heatmap displaying C-index values for all tested survival algorithms in training and validation cohorts (B) Scatter graph comparing model stability based on average C-index and standard deviation (C)

Dumbbell diagram visualizing cross-cohort generalization error of top-performing survival models (D–O) Kaplan–Meier survival curves, static ROC and time-dependent ROC plots for the RSF model in TCGA, GSE26901, GSE62254 and integrated pooled cohort (P–W) Kaplan–Meier survival stratification curves for eight separate external validation cohorts stratified by RSF risk scores (X) Violin plot showing the distribution of model-derived risk scores across all annotated single-cell lineages

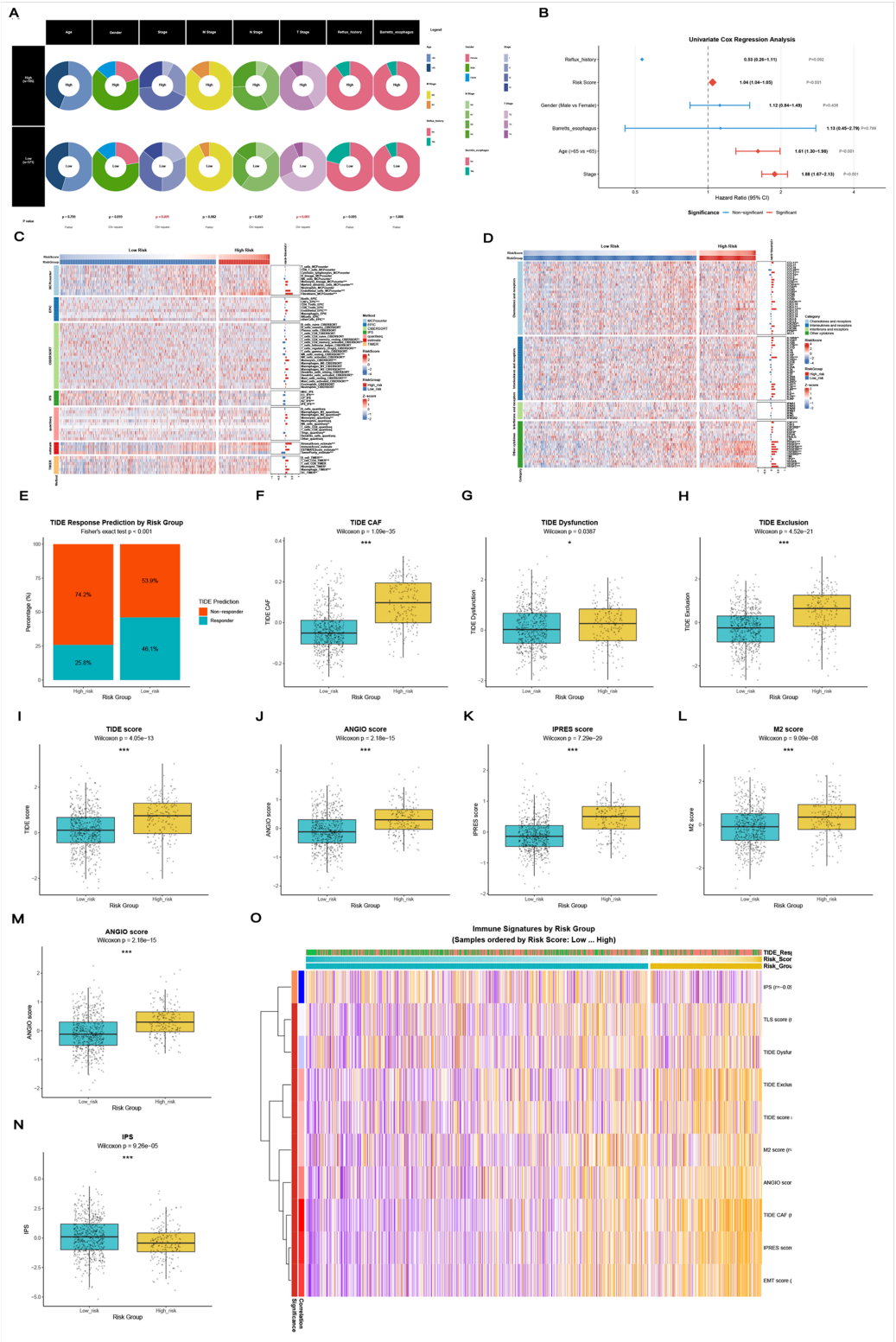


Figure 6

Clinical relevance, independent prognostic value, genomic features, immune microenvironment, and drug sensitivity of the risk score.

(A) Pie charts comparing the distribution of clinical covariates between high-risk and low-risk patient subgroups (B) Forest plot summarizing multivariate Cox regression outcomes for overall survival prediction (C) Rank-biserial correlation heatmap of immune cell abundances calculated by seven distinct deconvolution algorithms (D) Rank-biserial correlation heatmap covering chemokine and immune regulatory gene expression levels (E) Stacked bar graph showing the proportion of immunotherapy responders and non-responders predicted by TIDE algorithm (F–I) Box plots comparing TIDE_CAF, TIDE_Dysfunction, TIDE_Exclusion and overall TIDE scores across risk subgroups (J–M) Box plots of angiogenesis, IPRES and M2 macrophage signature scores in high/low-risk groups (N) Box plot of Immunophenoscore stratified by risk level (O) Correlation heatmap mapping the association between risk score and inflammatory/ stromal activation signatures. Supplementary Fig. 8 contains full pharmacogenomic susceptibility data. *p < 0.05, **p < 0.01, ***p < 0.01; Wilcoxon rank-sum test

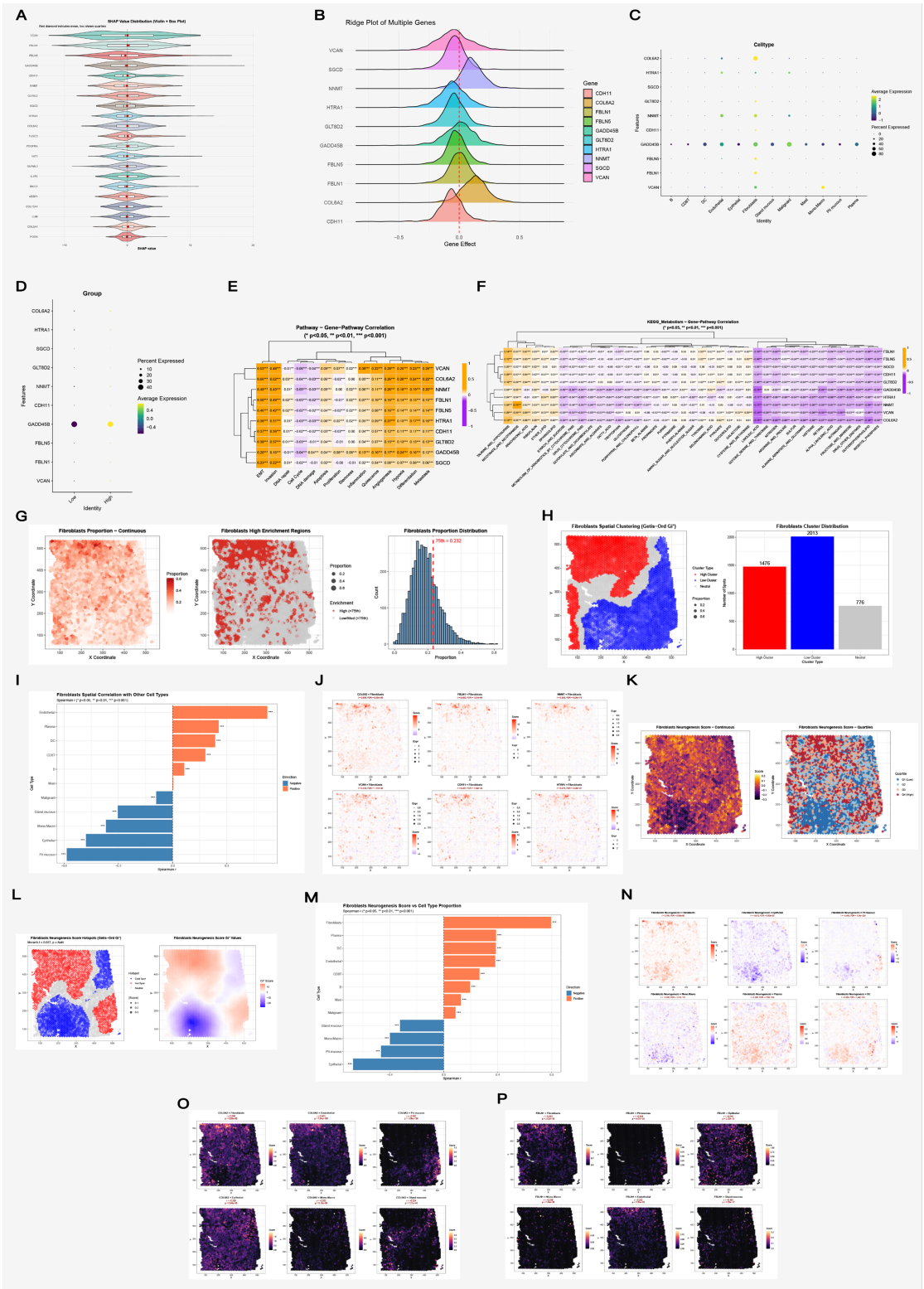


Figure 7

SHAP analysis identifies ten key prognostic genes with stromal-specific expression and spatial transcriptomic validation of fibroblast NM hotspots.

(A) Violin-box plot showing SHAP contribution values to rank all model input genes (B) Density ridge plots of DepMap CRISPR gene effect metrics for the ten core stromal genes (C) Dot diagram displaying

expression levels of core genes across twelve distinct cell lineages (D) Dot plot comparing core gene transcription between high- and low-neurogenesis mimicry microenvironments (E) Spearman correlation heatmap linking core gene expression to CancerSEA tumor functional signatures (F) Spearman correlation heatmap mapping core genes against KEGG metabolic pathway scores (G) Spatial distribution graph of fibroblast proportions with histogram of top 10% fibroblast-enriched tissue spots (H) Getis-Ord Gi* hotspot map reflecting clustered fibroblast distribution across tissue sections (I) Heatmap of spatial Spearman correlations between fibroblasts and other tissue cell subsets (J) Scatter plots showing spatial co-localization of representative core genes (VCAN, FBLN1, NNMT) with fibroblast-rich areas (K) Continuous and quartile-stratified spatial mapping of fibroblast neurogenesis mimicry scores (L) Gi*-based hotspot and coldspot layout of tissue neurogenesis mimicry signals (M) Spatial Spearman correlation heatmap between fibroblast NM scores and individual cell populations (N) Overlay visualization of NM transcriptional signals and cell-type proportion spatial patterns (O) Multi-lineage spatial co-localization panel for COL6A2 across six main cell groups (P) Multi-lineage spatial co-localization panel for FBLN1 across six main cell groups *p < 0.05, **p < 0.01

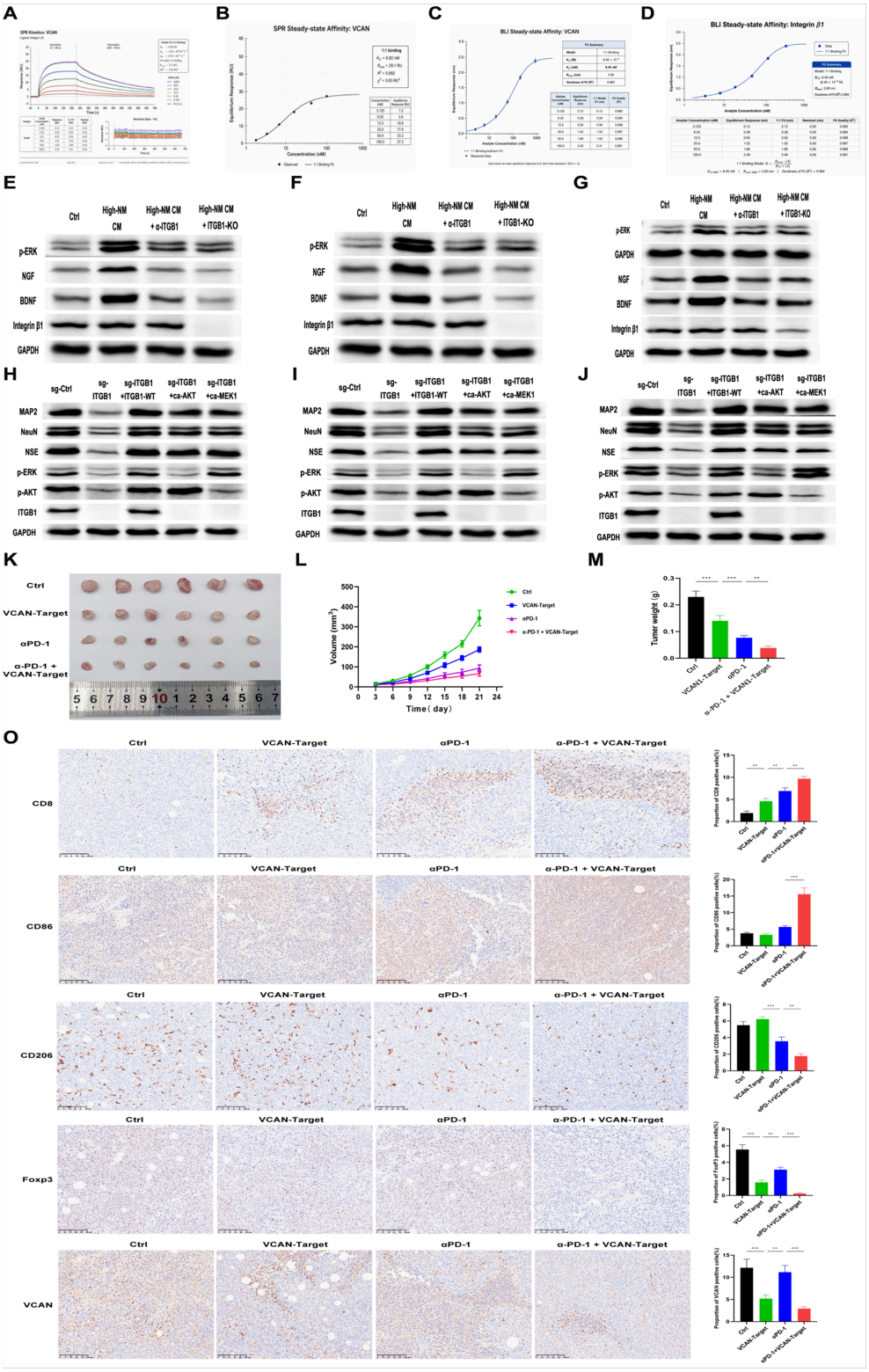


Figure 8

Biophysical binding of VCAN to Integrin β1 activates downstream AKT/ERK signaling to induce neurogenesis mimicry, and targeting VCAN sensitizes syngeneic orthotopic gastric tumors to α-PD-1 blockade.

(A) Surface plasmon resonance (SPR) kinetic binding sensorgrams of recombinant VCAN (3.125-100 nm) flowing over immobilized Integrin β1, with 1:1 binding fit curves (red lines) and residual plot. (B) SPR steady-state affinity saturation curve and Langmuir 1:1 binding model fitting for VCAN. (C, D) Bio-layer interferometry (BLI) steady-state equilibrium response curves and model fitting for VCAN (C) and Integrin β1 (D). (E–G) Western blot analysis of p-ERK, total ERK/GAPDH, NGF, BDNF, and Integrin β1 in AGS (E), MKN-45 (F), and HGC-27 (G) gastric cancer cells under four conditions: Control, High-NM CAF CM stimulation, CM with neutralizing α-ITGB1 antibody, and CM in CRISPR-mediated ITGB1-knockout (sg-ITGB1) cells. (H–J) Downstream signaling epistasis rescue assays in AGS (H), MKN-45 (I), and HGC-27 (J) cells. Western blot detection of neuronal markers (MAP2, NeuN, NSE), signaling activation (p-ERK, p-AKT), and ITGB1 in sg-Ctrl, sg-ITGB1, sg-ITGB1 + ITGB1-WT, sg-ITGB1 + constitutively active AKT (ca-AKT), and sg-ITGB1 + constitutively active MEK1 (ca-MEK1) cells. GAPDH served as the loading control. (K–M) In vivo therapeutic efficacy in an orthotopic syngeneic gastric cancer model (n = 6 mice/group) across four treatment arms: Control, VCAN-Target (sg-Vcan CAFs), α-PD-1 monotherapy, and α-PD-1 + VCAN-Target combination therapy. (K) Macroscopic images of excised orthotopic gastric tumors. (L) In vivo tumor volume growth kinetics measured over 21 days. (M) Statistical quantification of excised tumor weights at the endpoint.

Supplementary Files

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

- Supplementaryfile.docx
- SupplementaryFigure1.tif
- SupplementaryFigure2.tif
- SupplementaryFigure3.tif
- SupplementaryFigure4.tif
- SupplementaryFigure5.tif
- SupplementaryFigure6.tif
- SupplementaryFigure7.tif
- SupplementaryFigure8.tif
- SupplementaryFigure9.tif
- SupplementaryFigure10.tif
- SupplementaryFigure11.tif
- SupplementaryFigure12.tif
- SupplementaryFigure13.tif
- SupplementaryFigure14.tif
- SupplementaryFigure15.tif
- SupplementaryFigure16.tif

- SupplementaryFigure17.tif