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.

From Pan-Cancer Transcriptomics to Trial Alignment: Systematic Prioritization of PCSK9-Targeted Oncology Indications

peihong liang
School of Basic Medical Sciences, Guangzhou Medical University, China
huizhen ma
Shenzhen University Clinical Medical Academy Center
zongyi yin
yinzy@szu.edu.cn

Department of Hepatobiliary Surgery, Shenzhen University General Hospital, Shenzhen University

Research Article

Keywords: PCSK9, pan-cancer analysis, tumor immune microenvironment, single-cell transcriptomics, Notch pathway, molecular dynamics, hepatocellular carcinoma, immunotherapy response

Posted Date: September 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10699027/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

Proprotein convertase subtilisin/kexin type 9 (PCSK9) governs low-density lipoprotein receptor (LDLR) degradation, yet whether this cholesterol-regulating convertase intersects with tumor immunity remains incompletely defined. We integrated pan-cancer transcriptomics, single-cell and spatial transcriptomics, network biology, immune deconvolution, in silico perturbation, and structural drug validation to dissect PCSK9 function across tumors. Across 33 TCGA cancer types, PCSK9 was selectively up-regulated in digestive tract tumors (COAD log2FC = 2.71, FDR = 4.39 x 10^-18) and enriched for cholesterol metabolism, proliferation, and interferon-response programs. Immune deconvolution of 6,325 tumors revealed positive correlations with macrophages, Tregs, and MDSCs and negative correlations with CD8 + T cells, with PCSK9-high tumors showing an immune-excluded phenotype (TIDE rho = -0.46, P < 2.2 x 10^-16) and higher predicted immunotherapy response. Single-cell analysis of 71,914 cells from three HCC cohorts resolved PCSK9 as hepatocyte-restricted (18.54% positive) and tumor-up-regulated. PCSK9-high hepatocytes adopted an M2-dominant polarization, rewired intercellular communication, occupied a de-differentiated state, and revealed FGL1 as a core downstream node with dose-dependent Notch activation via a conserved NOTCH3 axis. Spatial transcriptomics confirmed PCSK9-macrophage co-localization and co-expression of FGL1, JAG2, and CD276. Drug screening and 100 ns molecular dynamics simulation nominated palbociclib (RMSD = 1.54 ± 0.25 A) as a direct PCSK9 binder and simvastatin as an indirect pathway modulator. Collectively, PCSK9 emerges as a metabolism-immunity crossover modulator whose tumor-cell-intrinsic expression orchestrates an immunosuppressive microenvironment, providing a multi-omics rationale for PCSK9-targeted therapy in digestive tract malignancies.

1. Introduction

Proprotein convertase subtilisin/kexin type 9 (PCSK9) first entered clinical view as a master regulator of familial hypercholesterolemia and has since become one of the most successful cardiovascular drug targets [1, 2]. By promoting lysosomal degradation of the low-density lipoprotein receptor (LDLR), PCSK9 raises circulating LDL cholesterol and drives atherosclerotic cardiovascular disease. Monoclonal antibodies (evolocumab, alirocumab) and an siRNA therapy (inclisiran) against PCSK9 have delivered robust, durable LDL lowering and cardiovascular risk reduction in large Phase III trials [1, 2]. Yet the very biology that makes PCSK9 indispensable in lipid metabolism–controlled receptor turnover at the cell surface–raises a provocative question: what does this convertase do when tumors exploit membrane-receptor and immune pathways for their own advantage?

In cancer, PCSK9 has been implicated in cholesterol reprogramming, proliferation, and immune evasion, but the direction of its effect is far from uniform. Experimental studies in colorectal and hepatocellular contexts point to pro-tumorigenic roles sustained by PI3K/AKT signaling and lipid metabolic rewiring [3, 4], whereas loss-of-function PCSK9 variants in melanoma models enhance CD8 + T-cell infiltration and improve anti-PD-1 responsiveness [5, 6]. Mechanistically, PCSK9 has been reported to dampen antitumor immunity by promoting lysosomal degradation of MHC class I, thereby attenuating antigen presentation

[7]. These observations are compelling individually, yet they remain siloed: the cellular origin of PCSK9, its downstream effector network, and its therapeutic implications within the tumor microenvironment (TME) have not been reconciled within a single, systematic framework.

Three gaps currently fragment the PCSK9-oncology field. First, most studies examine one or a few tumor types, so it remains unknown which cancer lineages harbor the most clinically consequential PCSK9 dysregulation. Second, analyses tend to emphasize either metabolism or immunity in isolation, obscuring the metabolic-immune crosstalk increasingly recognized as central to tumor progression. Third, no previous study has connected clinical prognosis, microenvironment composition, immunotherapy modelling, single-cell resolution, and structural drug validation within one pipeline. Key open questions therefore include: (i) which tumor types show the most clinically relevant PCSK9 dysregulation; (ii) which cell populations produce PCSK9 within the TME; (iii) what downstream regulatory networks link PCSK9 to immune suppression; and (iv) whether existing small-molecule scaffolds can be repurposed to target PCSK9 in oncology.

To bridge these gaps, we designed a comprehensive pan-cancer analysis that integrates nine analytical modules across 33 TCGA cancer types and three independent single-cell HCC cohorts. The workflow spans gene-function and protein-protein interaction (PPI) networks; pan-cancer expression profiling of PCSK9 and functionally related pathways; ceRNA regulatory network construction; genomic alteration analysis; tumor microenvironment and immune infiltration; single-cell transcriptomics with cell-cell communication; pseudotime trajectory and in silico perturbation; TIDE-based immunotherapy response modelling; and drug screening with molecular docking and molecular dynamics (MD) simulation. Spatial transcriptomics anchored these single-cell findings in tissue context.

The study makes five principal contributions: (i) it establishes PCSK9 as a hepatocyte-restricted, tumor-up-regulated gene operating at the cholesterol metabolism-immunity intersection; (ii) it shows that PCSK9-high tumor cells reshape the immune microenvironment through M2 polarization and rewired CellChat signaling; (iii) it identifies FGL1 and the NOTCH3-dedifferentiation axis as conserved downstream effectors using multi-method pseudotime and virtual perturbation; (iv) it provides spatial validation of PCSK9-macrophage co-localization and immunosuppressive marker co-expression; and (v) it nominates palbociclib and simvastatin as structurally informed PCSK9-targeting candidates through integrated drug screening and MD simulation.

2. Materials and Methods

This study followed a modular, omics-integrated workflow adapted from a prior indication-prioritization framework [10] using the integrated evaluation criteria detailed in Supplementary Table 5 and customized for PCSK9's metabolic-immune biology. Analyses were organized into nine core modules spanning bulk transcriptomics, network biology, immune deconvolution, single-cell genomics, spatial transcriptomics, and computational structural biology. All statistical analyses were performed in R v4.3.2 and Python v3.13 unless otherwise stated.

2.1 Data sources and preprocessing

Bulk RNA-seq expression and clinical annotations were obtained from The Cancer Genome Atlas (TCGA) across 33 cancer types (> 11,000 samples) using TCGAbiolinks (v2.25.2) and UCSC Xena Toil-derived matrices [11]. Tumor versus normal differential expression used paired and unpaired Wilcoxon rank-sum tests with Benjamini-Hochberg (BH) correction. Single-cell RNA-seq data comprised three independent cohorts: GSE125449 Set1 (HCC regrouped cohort, 9 samples, 3,667 cells, 868 hepatocytes), GSE125449 Set2 (intrahepatic cholangiocarcinoma [iCCA] regrouped cohort, 10 samples, 5,704 cells, 1,862 hepatocytes), and GSE149614 (10 paired HCC tumor/normal, 71,914 cells). Spatial transcriptomics used three 10x Visium HCC samples (Patient1: 1,025 spots; Patient2: 1,373 spots; Patient3: 755 spots). Batch effects were corrected with Harmony [33].

2.2 Pan-cancer differential expression and pathway analysis

PCSK9 expression across TCGA cancer types was evaluated with the R package TCGAplot using Wilcoxon rank-sum tests with BH correction. To capture the functional context beyond PCSK9 itself, we profiled the PCSK9 proprotein convertase family (PCSK1/2/4/5/6/7/9), Notch signaling core genes (NOTCH1-4, JAG1-2, DLL1/3/4, HES1, HEY1), and PPI cholesterol homeostasis genes (PCSK9, LDLR, APOB) in paired tumor-normal samples across 16 cancer types.

2.3 PPI network and functional enrichment

PCSK9-centered PPI networks were generated from STRING v12.0 (combined score > = 0.7) [15]. Gene Ontology (GO) biological process and KEGG enrichment were performed with clusterProfiler v4.8.0. Gene set enrichment analysis (GSEA) was conducted with gseapy using MSigDB Hallmark, KEGG 2021, and GO gene sets, with the ranked gene list generated from median log2 fold changes (tumor vs. normal) across 24 TCGA cancer types.

2.4 ceRNA regulatory network construction

Putative PCSK9-targeting microRNAs (miRNAs) were predicted using miRDB (score > = 80), miRWalk, and DIANA-microT; miRNAs supported by at least two databases were retained for robustness. For each candidate miRNA, predicted long non-coding RNA (lncRNA) targets were retrieved from ENCORI to build an mRNA-miRNA-lncRNA competing endogenous RNA (ceRNA) network. An integrated concentric network was constructed with PCSK9 at the center, PPI partners in the inner ring, miRNAs in the middle ring, and lncRNAs in the outer ring.

2.5 Tumor microenvironment and immune infiltration

Immune infiltration was estimated using four complementary algorithms–xCell [19], MCPcounter [20], CIBERSORT [21], and ESTIMATE [18]--chosen for their distinct methodological assumptions; associations consistent across algorithms were considered robust. Spearman correlation assessed associations between PCSK9 expression and immune cell abundances across 24 TCGA cancer types

(6,325 tumors). Immune-related functional genes were grouped into five categories: T-cell exhaustion markers, Treg signatures, M2 macrophage markers, angiogenesis factors, and EMT transcription factors, and PCSK9 correlations were evaluated cancer-type-specifically. To disentangle PCSK9 effects from canonical LDLR-cholesterol signaling, LDLR and APOB were included as control genes in lower gastrointestinal and liver cancer comparisons.

2.6 Genomic alteration analysis

PCSK9 mutation frequency, copy-number alteration, and alteration type distributions were summarized across TCGA cohorts using cBioPortal [22, 23].

2.7 TIDE immunotherapy response modelling

TIDE (Tumor Immune Dysfunction and Exclusion) scores were computed to estimate immunotherapy response potential by integrating T-cell dysfunction and exclusion components [17]. PCSK9-high versus PCSK9-low tumors (stratified by median expression) were compared for TIDE scores and predicted response rates across 24 cancer types. Because the composite score can mask complementary biology, TIDE subcomponents (Exclusion_Score, Dysfunction_Score) were additionally correlated with PCSK9 expression in the merged pan-cancer cohort (n = 6,368).

2.8 Single-cell transcriptomic analysis

scRNA-seq data were processed in Scanpy v1.9.6 [30] with UMAP dimensionality reduction (umap-learn v0.5.3 [37]) and Leiden clustering [31]. Cell types were annotated using canonical marker gene scores. For GSE149614, six major cell types were identified: Hepatocyte (20,782 cells), T/NK cell (25,591), Myeloid (15,947), B cell (3,685), Endothelial (3,644), and Fibroblast (2,266). PCSK9 expression was quantified per cell type and compared between tumor and adjacent normal tissues using Mann-Whitney U tests. Reflecting the sparse single-cell expression of PCSK9, PCSK9-high and PCSK9-low hepatocyte groups were defined using expression/non-expression dichotomization (HCC/iCCA cohorts) or top/bottom 25% stratification (GSE149614).

2.9 Cell-cell communication analysis

Ligand-receptor communication was inferred using CellChat [27] with the CellChatDB human ligand-receptor database. Differential outgoing communication patterns between PCSK9-high and PCSK9-low hepatocytes were quantified as percentage changes in communication probability for key pathways including MIF, SPP1, ApoE, VTN, FN1, and APP.

2.10 Pseudotime trajectory inference

To reduce single-method bias, three complementary methods were used for trajectory inference: Partition-based Graph Abstraction (PAGA) diffusion pseudotime [35], Monocle 3 [34], and scTour [36]. For all methods, the root cell was defined as the hepatocyte with highest PCSK9 expression, with trajectory direction defined from de-differentiated/stem-like to mature/differentiated states. Spearman

correlation between PCSK9 expression and pseudotime was computed across all three datasets and three methods to ensure cross-method consistency.

2.11 Virtual knock-out and over-expression analysis

Virtual knock-out (VKO) of PCSK9 was performed using scTenifoldKnk [29], which constructs a gene regulatory network from control cells and compares it against a perturbed network in which PCSK9 is virtually removed. Virtual over-expression (VOE) was performed using scTenifoldNet [29]. Differentially regulated genes satisfied adjusted $P < 0.05$. Because the 'FC' statistics from these tools represent network-perturbation effect sizes (Z-scores) rather than conventional expression fold changes, we interpreted them as measures of regulatory impact. Given known scalability limitations of scTenifoldNet CP-ALS tensor decomposition at full-gene scale, VOE analysis was restricted to the GSE149614 cohort using 5,000 highly variable genes; HCC and iCCA VOE results were excluded due to insufficient PCSK9-high cell numbers (HCC: 82; iCCA: 32) and algorithmic instability.

2.12 Notch pathway quantification

PCSK9-Notch associations were assessed across four orthogonal dimensions to guard against single-method bias: (1) DEG-level differences in Notch receptors, ligands, and effectors between PCSK9-high and PCSK9-low hepatocytes and between tumor and normal tissues; (2) subgroup Notch pathway scores (Receptor, Ligand, Effector) compared by Mann-Whitney U tests; (3) Spearman correlation between Notch gene expression and PAGA pseudotime; and (4) VKO/VOE Z-scores for Notch pathway genes. Results were integrated into a summary heatmap.

2.13 Scissor integration of bulk phenotypes

Scissor [28] was used to integrate bulk TCGA-LIHC phenotypes (overall survival [OS] and ImmuneScore) with single-cell expression data from the GSE125449 HCC and iCCA regrouped cohorts. Binomial regression identified Scissor+ cells associated with high immune infiltration in HCC, and Cox regression identified Scissor+ cells associated with poor prognosis in iCCA. PCSK9 expression was compared between Scissor+ cells and background using Wilcoxon tests.

2.14 Spatial transcriptomics analysis

Three HCC Visium samples were preprocessed and analyzed for spatial co-localization. Moran's I was computed to assess spatial autocorrelation of PCSK9-positive spots. Co-localization between PCSK9-positive spots and macrophage markers (CD68, CD163) was quantified by odds ratios (OR) with chi-square tests. Spearman correlations between PCSK9 and key single-cell molecules (FGL1, JAG2, ORM1, CD276, ALB, APOE, NOTCH3) were computed per patient and summarized across patients. Because Visium spots contain multiple cells, spatial correlation effect sizes are expected to be modest; we therefore emphasized cross-patient directional consistency when interpreting these results.

2.15 Drug screening, molecular docking, and molecular dynamics simulation

DrugReflector was used to screen the Connectivity Map (CMap) database (2,372 compounds) for agents predicted to reverse the PCSK9-high (malignant/hepatocyte-like) gene expression signature. Top candidates were ranked by logit scores and reversal probability. Molecular docking was performed with AutoDock Vina v1.2.3 [26] using the PCSK9 crystal structure (PDB: 2PMW) as receptor, with the docking box covering the catalytic-domain substrate-binding pocket. Candidate compounds included the DrugReflector top hits (BRD-K97168035, BRD-K31999759) and literature-reported PCSK9-associated compounds (olaparib, capmatinib, palbociclib, simvastatin). For each PCSK9-drug complex, a 100 ns all-atom molecular dynamics (MD) simulation was performed using GROMACS 2024.4 [38] on A800 80GB GPUs, following the EM -> NVT -> NPT -> production workflow (dt = 0.002 ps, 50,000,000 steps). Backbone root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), and protein-ligand close-contact counts (< 0.35 nm) were computed for Chain A of each complex.

2.16 Statistical analysis

Two-group comparisons used Wilcoxon or Mann-Whitney U tests; correlations used Spearman's rho with BH correction where indicated. All tests were two-sided with significance at $P < 0.05$ unless otherwise stated. Pseudotime and virtual perturbation analyses used rank-based statistics to handle sparse PCSK9 expression.

3. Results

3.1 PCSK9 is selectively up-regulated in digestive tract tumors and sits at the metabolism-immunity interface

To establish where and how PCSK9 is dysregulated across tumor lineages, we first performed paired differential expression analysis across 33 TCGA cancer types. PCSK9 was significantly up-regulated in multiple digestive cancers (Supplementary Fig. 1), including COAD (log2FC = 2.71, FDR = 4.39 x 10^-18), READ (log2FC = 2.65, FDR = 3.20 x 10^-5), ESCA (log2FC = 2.60, FDR = 5.10 x 10^-6), HNSC (log2FC = 2.47, FDR = 3.00 x 10^-17), STAD (log2FC = 1.75, FDR = 1.60 x 10^-8), and LIHC (log2FC = 1.12, FDR = 5.10 x 10^-10). In striking contrast, PCSK9 was down-regulated in KIRC (log2FC = -0.24, FDR = 7.40 x 10^-22) and KIRP (log2FC = -0.30, FDR = 3.90 x 10^-16). This pronounced bidirectional, lineage-specific pattern–enrichment in the digestive tract, depletion in the kidney–immediately frames PCSK9 not as a generic tumor marker but as a tissue-context-dependent modulator (Fig. 2a).

Analysis of the broader proprotein convertase family (PCSK1/2/4/5/6/7/9) revealed striking functional diversification. PCSK9 displayed gastrointestinal tumor-specific high expression, whereas PCSK5 and PCSK6 were broadly up-regulated across most cancer types, and PCSK1 was enriched in THCA, consistent with its neuroendocrine hormone-processing function (Fig. 2b). The Notch signaling pathway–a core developmental program governing cell fate–was globally activated in digestive tumors, with NOTCH1 up-regulated in COAD, STAD, and ESCA and JAG2 significantly elevated in LIHC and HNSC

(Fig. 2c). Finally, the PPI cholesterol homeostasis pathway showed a complementary architecture: PCSK9 up-regulation coincided with LDLR down-regulation in digestive tumors, consistent with active PCSK9-mediated LDLR degradation, while APOB was specifically up-regulated in LIHC, mirroring the hepatic lipid dysregulation characteristic of liver cancer (Fig. 2d). The parallel activation of PCSK9 and Notch signaling in the same digestive lineages prompted us to examine their functional relationship at single-cell resolution (see below).

STRING-based PPI analysis placed PCSK9 at the center of a lipid-metabolism interactome, with direct interactions with LDLR, VLDLR, APOB, HMGCR, and GOLPH3 (Fig. 1a). GSEA of PCSK9-associated genes using KEGG pathways revealed positive enrichment in Cell Cycle (NES = 2.10, FDR = 0.003) and Bladder Cancer (NES = 2.04, FDR = 0.003), alongside negative enrichment in metabolic pathways including Beta-Alanine Metabolism (NES = -2.29) and Butanoate Metabolism (NES = -2.87) (Fig. 1b). Hallmark pathway GSEA further confirmed proliferation-associated enrichment (E2F Targets: NES = 2.54, FDR < 0.001; G2M Checkpoint: NES = 2.32, FDR < 0.001; MYC Targets: NES = 1.81–1.96, FDR < 0.001) and immune signaling (Interferon Alpha/Gamma Response: NES = 1.66–1.80, FDR < 0.001), while metabolic differentiation pathways were negatively enriched (Bile Acid Metabolism: NES = -2.13; Fatty Acid Metabolism: NES = -1.75) (Fig. 1c). Collectively, these results depict PCSK9 as a metabolic switch: its expression re-routes cells away from lipid differentiation toward proliferation and immune signaling, a phenotype that is directly interrogated in the single-cell data below.

To dissect post-transcriptional regulation, we integrated miRNA predictions from miRDB, miRWalk, and ENCORI, identifying the top 20 miRNAs targeting PCSK9 (Fig. 1d). An integrated concentric ceRNA network placed PCSK9 at the center, with PPI partners in the inner ring, miRNAs in the middle ring, and lncRNAs predicted to sponge these miRNAs in the outer ring (Fig. 1e). This layered regulatory architecture suggests that PCSK9 expression is shaped by ceRNA-level control, adding a post-transcriptional dimension to its metabolic and immune functions that may ultimately be exploitable for therapeutic modulation.

3.2 PCSK9 shapes an immunosuppressive, immune-excluded tumor microenvironment and predicts immunotherapy response

Having established digestive-specific up-regulation, we next asked whether PCSK9 co-varies with the immune landscape that determines immunotherapy outcomes. Integrating xCell, MCPcounter, CIBERSORT, and ESTIMATE across 24 TCGA cancer types (6,325 tumors; Supplementary Table 1), we found a remarkably consistent immunoregulatory pattern: PCSK9 was positively correlated with tumor-associated macrophages, Tregs, MDSCs, and cancer-associated fibroblasts (Supplementary Table 4), with the strongest associations in STAD (rho = 0.71–0.75), HNSC (rho = 0.66–0.67), LIHC (rho = 0.67), BLCA (rho = 0.65–0.74), and GBM (rho = 0.63–0.71) (all FDR < 1 x 10^-40). Conversely, PCSK9 tracked negatively with CD8 + effector T cells and exhausted T cells, most strongly in TGCT (rho = -0.71, P < 1 x 10^-23) (Fig. 3a). The convergence of four independent algorithms–each with distinct assumptions–on

the same signature strengthened our confidence that this association is biological rather than algorithmic.

CIBERSORT independently confirmed the Treg association, with PCSK9 positively correlated with Treg infiltration in 18 cancer types (STAD, HNSC, LIHC, and BLCA all rho > 0.6), while M2 macrophage associations showed cancer-type heterogeneity (Fig. 3b); cancer-specific PCSK9-CD8 + T-cell correlations are detailed in Supplementary Table 2. At the molecular level, PCSK9 was positively associated with Treg core markers (FOXP3, IL2RA, CTLA4), M2 macrophage markers (CD163, MRC1, MS4A4A), and EMT transcription factors (SNAI1, SNAI2, TWIST1, VIM, CDH2) in BLCA, STAD, and HNSC, suggesting that PCSK9 may both recruit immunosuppressive cells and promote epithelial-mesenchymal transition within the same tumors (Fig. 3c). Significance analysis across the pan-cancer panel confirmed that PCSK9 correlations with macrophages, Tregs, MDSCs, and CD8 + T cells reached FDR < 0.05 in over two-thirds of cancer types, attesting to a broad, not incidental, immunomodulatory program (Fig. 3d).

To determine whether these immune associations merely reflect PCSK9's canonical LDLR-cholesterol function, we benchmarked PCSK9 against LDLR and APOB in COAD, READ, and LIHC. LDLR showed a concordant but markedly weaker correlation pattern, whereas APOB displayed virtually no significant immune associations (Fig. 3e). This dissociation is a key negative control: it indicates that PCSK9's immunomodulatory effects are partially independent of its classical LDLR-degradation function, consistent with non-canonical, immune-specific activities. Representative scatter plots in four digestive cancers confirmed strong PCSK9-immunosuppressive-cell correlations–LIHC vs MDSC (R = 0.80, n = 421), COAD vs Treg (R = 0.80, n = 522), READ vs macrophages (R = 0.81, n = 176), and STAD vs Treg (R = 0.84, n = 448), all P < 0.001 (Fig. 3f).

Given this immunosuppressive association, we asked whether PCSK9 informs immune checkpoint inhibitor (ICI) response using TIDE. In the merged pan-cancer cohort (n = 6,368), PCSK9 correlated negatively with TIDE scores (Spearman rho = -0.46, P < 2.2 x 10^-16), suggesting that PCSK9-high tumors may be more responsive to immunotherapy (Fig. 3g). Decomposing TIDE subcomponents resolved a seemingly paradoxical picture: PCSK9 correlated strongly and positively with immune-exclusion scores (Exclusion_Score: rho = 0.39, P < 1 x 10^-226) but only weakly with T-cell dysfunction (Dysfunction_Score: rho = -0.03). Together with the Treg/MDSC enrichment, this points to an immune-excluded phenotype–effector CD8 + T cells barred from the tumor parenchyma by an immunosuppressive stromal barrier–rather than a T-cell-inflamed, dysfunctional state. Such exclusion phenotypes typically respond better to ICI than dysfunctional/inflamed tumors, reconciling the apparent contradiction between immunosuppressive enrichment and improved predicted response. The pattern held in an independent TIDE immunotherapy cohort (Fig. 3h), and PCSK9-high tumors showed higher predicted response rates in 16 of 24 cancer types, including KIRC, KIRP, HNSC, GBM, ESCA, PAAD, and STAD (Fig. 3i).

3.3 Single-cell resolution identifies hepatocyte-restricted PCSK9 that reshapes the immune microenvironment

To translate our bulk observations into cell-level mechanism, we needed to know which cells actually express PCSK9 and how they behave. Integrating three independent scRNA-seq cohorts–GSE125449 Set1 (HCC, 3,667 cells), GSE125449 Set2 (iCCA, 5,704 cells), and GSE149614 (paired HCC tumor/normal, 71,914 cells)--we first resolved the cellular landscape of GSE149614, yielding six major cell types after unified QC and batch correction: Hepatocyte (20,782), T/NK cell (25,591), Myeloid (15,947), B cell (3,685), Endothelial (3,644), and Fibroblast (2,266), with clean UMAP separation (Fig. 4a).

PCSK9 expression proved strikingly cell-type-specific: hepatocytes showed the highest expression rate (18.54% positive, mean = 0.090), far exceeding immune cells (T/NK 1.01%, Myeloid 1.03%, B cell 2.39%) and stromal cells (Endothelial 0.41%, Fibroblast 0.49%) (Fig. 4b). Paired tumor-normal comparison in GSE149614 demonstrated significant PCSK9 up-regulation in tumor tissue (9.81% positive vs 0.41%, mean 0.047 vs 0.003, Mann-Whitney $P < 2.2 \times 10^{-16}$) (Fig. 4c). This hepatocyte-restricted, tumor-elevated pattern resolves the cellular-origin question raised in the Introduction and confirms that PCSK9 acts primarily from malignant hepatocytes, consistent with our pan-cancer GSEA findings.

To characterize the functional state of PCSK9-high hepatocytes, we stratified hepatocytes into PCSK9-high and PCSK9-low/no groups and scored nine immune gene sets. In HCC, PCSK9-high hepatocytes showed significantly elevated cholesterol metabolism (mean 1.228 vs 0.896, $P = 1.90 \times 10^{-14}$), M2 polarization (0.195 vs 0.123, $P = 1.62 \times 10^{-5}$), Treg (0.021 vs 0.016, $P = 1.49 \times 10^{-3}$), and Cytotoxic (0.053 vs 0.035, $P = 7.44 \times 10^{-3}$) scores; in iCCA, cholesterol metabolism (0.536 vs 0.255, $P = 3.72 \times 10^{-8}$), Treg (0.024 vs 0.012, $P = 2.51 \times 10^{-6}$), and MHCII (0.781 vs 0.631, $P = 3.25 \times 10^{-2}$) scores were elevated (Fig. 4d). Critically, the M1-M2 score difference was significantly lower in PCSK9-high hepatocytes (-0.191 vs -0.118), indicating an M2-dominant polarization state that mirrors the pan-cancer PCSK9-M2 correlation (Fig. 4e).

CellChat analysis in GSE149614 revealed that PCSK9-high hepatocytes actively rewired communication with the surrounding immune milieu. Pro-inflammatory signals MIF (to Myeloid: -32.2%, to T/NK: -32.3%) and SPP1 (to T/NK: -67.4%, to Myeloid: -67.5%) were markedly attenuated, whereas lipid-metabolism signals ApoE to Myeloid (+ 20.4%), VTN to Myeloid (+ 12.8%), and APP to Myeloid (+ 337.5%) were enhanced (Fig. 4f). A chord diagram centered on PCSK9-high hepatocytes illustrated this reconfiguration (Fig. 4g). The convergent logic is instructive: PCSK9-high hepatocytes suppress pro-inflammatory, pro-phagocytic cues (MIF, SPP1) while amplifying lipid-metabolism and ECM signals (ApoE, VTN, FN1) directed at myeloid cells–an elegant molecular handshake by which metabolic reprogramming is translated into an immunosuppressive cellular message, echoing the GSEA-identified cholesterol program.

3.4 PCSK9-high hepatocytes are de-differentiated, activate Notch signaling, and converge on FGL1

To determine the developmental state of PCSK9-high hepatocytes, we performed pseudotime trajectory inference with three complementary methods. PAGA diffusion pseudotime revealed that PCSK9 expression was negatively correlated with pseudotime in all three datasets (GSE149614: rho = -0.090, P

= 8.88 x 10^-39; HCC: rho = -0.134, P = 7.25 x 10^-5; iCCA: rho = -0.127, P = 3.50 x 10^-8), positioning PCSK9-high cells at the trajectory origin, i.e., a de-differentiated/stem-like state (Fig. 5a). Monocle 3 independently corroborated this direction (GSE149614: rho = -0.029, P = 3.47 x 10^-5; HCC: rho = -0.107, P = 1.52 x 10^-3; iCCA: rho = -0.156, P = 1.15 x 10^-11) (Fig. 5b), and scTour provided a third, concordant method. In GSE149614, mature hepatocyte markers (ALB, ARG1, APOA1, APOE, CD74) increased with pseudotime, whereas Notch pathway genes (HES1, HEY1, JAG1, NOTCH3) decreased (Fig. 5c), indicating that PCSK9-high cells occupy a de-differentiated state with active Notch signaling–consistent with the digestive-tumor Notch activation observed in bulk data.

To move beyond association toward network-level inference, we performed virtual knock-out (VKO) of PCSK9 using scTenifoldKnk in GSE149614. Five genes reached adjusted P < 0.05: ORM1 (FC = 215.36, padj = 2.33 x 10^-46), FGL1 (FC = 173.98, padj = 1.67 x 10^-37), ALB (FC = 15.58, padj = 9.90 x 10^-3), and GC (FC = 14.17, padj = 1.67 x 10^-2) (Fig. 5d). Virtual over-expression (VOE) with scTenifoldNet identified 17 significant genes, including ALB (FC = 161.12), HLA-DPB1 (FC = 42.57, padj = 1.71 x 10^-8), ANGPT2 (FC = 39.03), and FGL1 (FC = 11.92, padj = 2.83 x 10^-2) (Fig. 5e). Notably, these 'FC' values represent network-perturbation effect sizes (Z-scores) reflecting the regulatory impact of PCSK9 perturbation, not conventional expression fold changes. FGL1 was significant in both perturbations (VKO: FC = 174; VOE: FC = 12), emerging as the most robustly PCSK9-linked immune candidate. As a well-characterized high-affinity LAG-3 ligand that suppresses T-cell activation, FGL1 provides a direct molecular bridge from hepatocyte PCSK9 to immune suppression. The HLA-DPB1 up-regulation observed upon VOE further suggests that PCSK9 may modulate antigen-presentation programs, consistent with its pan-cancer immunotherapy signature.

Because Notch signaling was co-activated with PCSK9 in bulk digestive tumors and conversely correlated with pseudotime in PCSK9-high cells, we tested the PCSK9-Notch relationship across four orthogonal dimensions (Fig. 5f). (1) At the DEG level, Notch receptors/ligands/effectors were broadly up-regulated in GSE149614 tumor versus normal, with NOTCH3, HES1, and JAG1 differentially expressed in PCSK9-high hepatocytes. (2) By subgroup scoring, PCSK9-high hepatocytes showed significantly elevated Notch Receptor (P = 1.8 x 10^-4), Effector (P = 2.8 x 10^-3), and Ligand (P = 3.7 x 10^-2) scores in GSE149614. (3) By pseudotime correlation, NOTCH3 was negatively correlated with PAGA pseudotime in both HCC (rho = -0.186, P = 3.7 x 10^-8) and iCCA, as was HES1, linking Notch activation to de-differentiation (Fig. 5g). (4) By virtual perturbation, VKO reduced Notch gene Z-scores (i.e., PCSK9 normally restrains Notch), while VOE increased DLL1 (Z = + 2.58), NOTCH3 (Z = + 1.97), JAG1 (Z = + 1.48), and HES1 (Z = + 1.25). Integration of these four dimensions revealed dose-dependent Notch activation by PCSK9, with the NOTCH3-dedifferentiation axis conserved across all three datasets as the core downstream effector.

To connect these single-cell programs to bulk patient phenotypes, we applied Scissor using TCGA-LIHC ImmuneScore (HCC cohort) and OS (iCCA cohort). In HCC, binomial regression identified 927 Scissor+ cells associated with high immune infiltration (73.5% tumor-associated macrophages) and significantly elevated PCSK9 (mean 0.037 vs 0.007, Wilcoxon P = 2.78 x 10^-9). In iCCA, Cox regression identified 237

Scissor+ cells associated with poor prognosis (85.2% malignant cells), also with elevated PCSK9 (mean 0.021 vs 0.004, P = 4.33 x 10^-10) (Fig. 5h-J). Thus, from an orthogonal bulk-phenotype-mapping standpoint, PCSK9-high cells are concordantly linked to immune infiltration and adverse prognosis, closing the loop with the pan-cancer findings.

Spatial transcriptomics in three HCC Visium samples (Supplementary Figs. 2 and 3) anchored these findings in tissue. PCSK9-positive spots exhibited significant positive spatial autocorrelation (Moran's I: Patient1 = 0.058, Patient2 = 0.026, Patient3 = 0.163), indicating that PCSK9-high regions form contiguous functional domains rather than random scatter. PCSK9-positive spots co-localized with CD68 + macrophages (Patient1: OR = 1.68, P = 6.86 x 10^-4; Patient3: OR = 1.59, P = 1.54 x 10^-4) and CD163 + macrophages (Patient2: OR = 1.34, P = 7.20 x 10^-3) (Fig. 5k). Key single-cell molecules showed directionally consistent spatial co-expression with PCSK9: FGL1 (mean rho = 0.050), JAG2 (mean rho = 0.065), ORM1 (mean rho = 0.058), and CD276/B7-H3 (mean rho = 0.048), each consistent across 3/3 patients, whereas mature hepatocyte markers ALB/APOE and NOTCH3 showed weak or inconsistent correlations (Fig. 5l). PCSK9-high regions displayed reduced mature liver function (ALB, APOE, Liver_secreted scores) and elevated JAG2/CD276, mirroring the de-differentiated, immunosuppressive phenotype defined at single-cell resolution. This in situ convergence–PCSK9, FGL1, JAG2, and CD276 co-occurring with macrophage infiltration–substantially strengthens the biological plausibility of the proposed PCSK9-Notch-FGL1 axis.

3.5 Drug screening and structural simulation nominate palbociclib and simvastatin as PCSK9-targeting candidates

Finally, to identify clinically actionable strategies, we asked whether existing compounds could reverse the PCSK9-high malignant phenotype. DrugReflector screening of the CMap database (2,372 compounds) ranked BRD-K97168035 (logit = 9.62, prob = 0.047) and BRD-K31999759 (logit = 9.21, prob = 0.031) first and second, both sulfonamide small molecules sharing a piperazine core scaffold (Fig. 6a).

To complement this transcriptional logic with structural evidence, we performed AutoDock Vina docking against the PCSK9 catalytic domain (PDB: 2PMW) for six candidates. Olaparib showed the strongest affinity (Vina score = -8.69 kcal/mol), followed by capmatinib (-8.01 kcal/mol), palbociclib (-7.36 kcal/mol), and simvastatin (-6.39 kcal/mol) (Fig. 6b-G), with binding poses engaging key catalytic residues.

Because static docking cannot distinguish a stable binder from a transiently docked pose, we subjected all six complexes to 100 ns all-atom MD simulations. Backbone C-alpha RMSD of Chain A revealed a clear stability hierarchy: the palbociclib complex showed the best scaffold stability (RMSD = 1.54 ± 0.25 A, Rg = 2.56 ± 0.02 nm), followed by olaparib (RMSD = 4.37 ± 8.78 A) and capmatinib (RMSD = 3.94 ± 8.40 A), though with greater fluctuation. The two DrugReflector hits showed high RMSD (BRD-K31999759: 31.56 ± 5.81 A; BRD-K97168035: 24.94 ± 10.32 A), indicating that their predicted modes require optimization. Simvastatin showed RMSD = 22.38 ± 21.24 A with Rg = 3.06 ± 0.50 nm and 15.9 ± 14.7 close contacts, consistent with conformational change or pocket dissociation during the unconstrained

production phase. Protein-ligand close-contact analysis showed that palbociclib (27.6 ± 11.7) and BRD-K31999759 (28.2 ± 8.4) maintained the most persistent interactions (Fig. 6h-l).

Integrating screening, docking, and MD evidence yields a mechanistically nuanced picture. Palbociclib performed well in both docking affinity (-7.36 kcal/mol) and MD stability (RMSD = 1.54 ± 0.25 A); together with its CDK4/6 inhibition, this positions it as the strongest direct PCSK9-binding candidate, although its on-target PCSK9 engagement requires orthogonal experimental validation. Simvastatin, an approved HMG-CoA reductase inhibitor, operates through a distinct indirect route: by reducing intracellular cholesterol synthesis it transcriptionally up-regulates LDLR, indirectly antagonizing PCSK9-mediated LDLR degradation. Although simvastatin did not maintain a stable direct pose in our MD simulations (RMSD = 22.38 ± 21.24 A)--consistent with its indirect mode of action–its established clinical safety and epidemiological signal for cancer-risk reduction warrant further structural optimization of PCSK9 binding.

4. Discussion

This integrated pan-cancer study repositions PCSK9 from a cardiovascular lipid regulator to a context-dependent metabolism-immunity crossover modulator in oncology. By synthesizing expression, network, immune, single-cell, perturbation, spatial, and structural evidence within a single framework, we demonstrate that hepatocyte-derived PCSK9 orchestrates immunosuppressive microenvironments through coordinated metabolic reprogramming, Notch pathway activation, and rewired intercellular communication.

The central advance is the demonstration that PCSK9 associations diverge systematically across tumor lineages while nevertheless converging on a common metabolic-immune mechanism. Gastrointestinal and upper aerodigestive cancers showed the strongest PCSK9 up-regulation, enriched for cholesterol metabolism and proliferation programs, whereas renal cancers showed significant down-regulation [9]. This directionality aligns with tissue-specific cholesterol handling: digestive tumors actively invoke PCSK9-mediated LDLR degradation to meet proliferative lipid demands, whereas renal tumors appear to suppress cholesterol biosynthetic flux through alternative routes. The GSEA-identified shift from metabolic differentiation (Bile Acid, Fatty Acid Metabolism) toward proliferation (E2F, G2M, MYC) and immune signaling (Interferon Alpha/Gamma) casts PCSK9 as a metabolic switch that tumors exploit for both growth and immune evasion.

A key finding is the partial independence of PCSK9's immune functions from its canonical LDLR pathway. LDLR showed concordant but weaker immune correlations, whereas APOB displayed essentially no immune associations, indicating that PCSK9's immunomodulatory effects cannot be fully attributed to cholesterol-lowering mechanisms. This is consistent with evidence that PCSK9 can suppress MHC class I antigen presentation through direct protein-protein interaction [7], and it extends that mechanism to a pan-cancer context. The TIDE analysis reconciles an apparent paradox: PCSK9-high tumors simultaneously enrich immunosuppressive cells and show higher predicted ICI response [8]. The

resolution lies in phenotype, not quantity–PCSK9-high tumors are immune-excluded rather than dysfunctional, and because exclusion is reversible by disrupting the stromal barrier, these tumors are more amenable to ICI than the dysfunctional/inflamed phenotype [12].

Single-cell analysis supplied the cellular resolution missing from bulk deconvolution. PCSK9 was confirmed as hepatocyte-restricted (18.54% positive) and tumor-up-regulated (9.81% vs 0.41%), resolving the question of cellular origin. PCSK9-high hepatocytes adopted an M2-dominant polarization and rewired CellChat communication–attenuating pro-inflammatory MIF/SPP1 while amplifying lipid-metabolism (ApoE, VTN) and ECM (FN1) signals. This dual mechanism–suppressing inflammation while supplying immunosuppressive metabolic cues–provides a coherent molecular explanation for the pan-cancer PCSK9-macrophage/Treg correlations. Spatial validation of PCSK9-CD68 + macrophage co-localization (OR = 1.59–1.68) confirms that these interactions occur in situ rather than as dissociated scRNA-seq artifacts.

Virtual perturbation and pseudotime analysis identified FGL1 as a core PCSK9 downstream node, significant in both VKO (FC = 174) and VOE (FC = 12). FGL1 is a hepatocyte-derived immune-checkpoint ligand that binds LAG-3 to suppress T-cell activation, providing a direct mechanistic link between PCSK9-high hepatocytes and T-cell exclusion–and, notably, operating independently of the PD-L1/PD-1 axis, which may have implications for checkpoint-combination strategies. The dose-dependent Notch activation, with the NOTCH3-dedifferentiation axis conserved across three datasets and three pseudotime methods, positions Notch signaling as a key PCSK9 effector. This is consistent with Notch's established role in hepatocyte de-differentiation and tumor stemness, and it suggests the PCSK9-Notch axis as a candidate therapeutic intervention point.

The drug-screening and structural pipeline nominated palbociclib and simvastatin through complementary mechanisms. Palbociclib's excellent MD stability (RMSD = 1.54 ± 0.25 A) and CDK4/6 inhibition offer a dual benefit of direct PCSK9 binding and antagonism of the proliferative programs enriched in PCSK9-high tumors. Simvastatin provides an indirect LDLR-restoration route via HMG-CoA reductase inhibition, with established clinical safety supporting relatively rapid translation. The BRD compounds, despite high DrugReflector scores, showed suboptimal MD stability–underscoring the importance of dynamic validation beyond static docking and illustrating the value of integrating transcriptional and structural evidence.

Several limitations should temper interpretation. Most analyses are correlative and derived from bulk TCGA data with platform heterogeneity; multivariate Cox survival analyses are detailed in Supplementary Table 3; deconvolution algorithms provide approximate immune abundances and may conflate tumor-cell-intrinsic PCSK9 with stromal signals. Single-cell and spatial analyses centered on HCC/iCCA cohorts, so generalization to non-hepatic tumors requires further validation. The VKO/VOE results are computational predictions requiring experimental confirmation, and iCCA VOE was excluded due to algorithmic limitations with sparse PCSK9-high cells. MD simulations used a single PCSK9 conformation and a 100 ns timescale, which may not capture all binding modes or slow conformational transitions; in

addition, the 'FC' statistics from scTenifoldKnk/Net reflect network-perturbation effect sizes rather than expression changes, and Visium spot resolution imposes modest spatial correlation effect sizes. These considerations position our mechanistic claims as hypothesis-generating, with causality to be established by functional experiments.

Future studies should prioritize (i) experimental validation of FGL1 and NOTCH3 as PCSK9 effectors using CRISPR perturbation in hepatic organoids; (ii) spatial multi-omics mapping of PCSK9, FGL1, and immune cells across tumor types beyond HCC; (iii) prospective biomarker-enriched trial designs exploiting PCSK9-TIDE joint stratification; and (iv) structure-based optimization of the simvastatin and BRD scaffolds to enhance PCSK9 binding affinity and MD stability.

5. Conclusions

PCSK9 functions as a metabolism-immunity crossover modulator whose pan-cancer associations are highly context-dependent. Omics-integrated evidence from pan-cancer transcriptomics, single-cell and spatial genomics, virtual perturbation, and structural drug validation converges on a unified mechanism: hepatocyte-derived, tumor-up-regulated PCSK9 couples cholesterol-metabolism reprogramming with M2/Treg immunosuppression, reshapes intercellular communication through metabolic signal rewiring, drives hepatocyte de-differentiation via the NOTCH3 axis, and predicts immunotherapy response through an immune-excluded phenotype. FGL1 emerges as a core downstream node, while palbociclib (direct) and simvastatin (indirect) are nominated as structurally informed therapeutic candidates. These findings establish PCSK9 as a tractable, metabolism-immunity target for translational oncology, particularly in digestive tract tumors and HCC.

Abbreviations

BH, Benjamini-Hochberg; BLCA, bladder urothelial carcinoma; CAF, cancer-associated fibroblast; CMap, Connectivity Map; CD8+ T, CD8-positive T cell; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; Cox, Cox proportional hazards; CTRP, Cancer Therapeutics Response Portal; DEG, differentially expressed gene; ECM, extracellular matrix; EMT, epithelial-mesenchymal transition; ESCA, esophageal carcinoma; FDR, false discovery rate; GBM, glioblastoma multiforme; GSEA, gene set enrichment analysis; GRO, GROMACS; HCC, hepatocellular carcinoma; HLA, human leukocyte antigen; HMGCR, 3-hydroxy-3-methylglutaryl-CoA reductase; HNSC, head and neck squamous cell carcinoma; iCCA, intrahepatic cholangiocarcinoma; ICI, immune checkpoint inhibitor; KEGG, Kyoto Encyclopedia of Genes and Genomes; KIRC, kidney renal clear cell carcinoma; LAG-3, lymphocyte-activation gene 3; LIHC, liver hepatocellular carcinoma; LDLR, low-density lipoprotein receptor; MD, molecular dynamics; MDSC, myeloid-derived suppressor cell; MCPcounter, Microenvironment Cell Populations counter; MHC, major histocompatibility complex; MSI, microsatellite instability; NPT, isothermal-isobaric ensemble; NVT, canonical ensemble; OR, odds ratio; OS, overall survival; PAAD, pancreatic adenocarcinoma; PAGA, Partition-based Graph Abstraction; PDB, Protein Data Bank; PRAD, prostate adenocarcinoma; PPI, protein-protein interaction; QC, quality control; READ, rectum adenocarcinoma; RMSD, root-mean-square

deviation; RMSF, root-mean-square fluctuation; Rg, radius of gyration; SASA, solvent-accessible surface area; scRNA-seq, single-cell RNA sequencing; STAD, stomach adenocarcinoma; TCGA, The Cancer Genome Atlas; TGCT, testicular germ cell tumors; THCA, thyroid carcinoma; TIDE, Tumor Immune Dysfunction and Exclusion; TMB, tumor mutational burden; TME, tumor microenvironment; TPM, transcripts per million; Treg, regulatory T cell; UMAP, Uniform Manifold Approximation and Projection; VKO, virtual knock-out; VOE, virtual over-expression

Declarations

Funding: This study was supported by grants from the National Natural Science Foundation of China (82104597), Hubei Chen Xiaoping Science and Technology Development Foundation (CXPJJH122002-068), Shenzhen Medical Research Fund (C2401014), and HaiYa Young Scientist Foundation of Shenzhen University General Hospital (2025-HY013). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions: Z.Y. (Zongyi Yin), Ph.L. (Peihong Liang): conceptualization, methodology, formal analysis, investigation, data curation, visualization, writing—original draft. Huizhen Ma and Huilin Ma: data curation, validation, and investigation. Xiaowu Li: resources and clinical data curation. Zongyi Yin and Dongying Chen: supervision, funding acquisition, project administration, writing—review and editing. All authors reviewed and approved the final manuscript.

Competing interests: The authors declare no competing interests.

Data availability: Bulk RNA-seq expression and clinical data were obtained from The Cancer Genome Atlas (TCGA) via the following project-level sources: TCGAbiolinks R package (https://bioconductor.org/packages/TCGAbiolinks/) and UCSC Xena Toil-derived matrices (https://toil.ucsc.edu/). Individual TCGA project/cohort codes analyzed in this study: TCGA-LIHC, TCGA-COAD, TCGA-READ, TCGA-ESCA, TCGA-STAD, TCGA-HNSC, TCGA-KIRC, TCGA-KIRP, TCGA-THCA, TCGA-PAAD, TCGA-GBM, TCGA-BLCA, TCGA-UVM, TCGA-SARC, TCGA-SKCM, TCGA-TGCT, TCGA-LUSC, TCGA-BRCA, TCGA-PRAD, TCGA-THYM, TCGA-OV, TCGA-UCS, TCGA-CESC, TCGA-ACCF, TCGA-KICH, TCGA-LAML, TCGA-MESO, TCGA-DEL, TCGA-PCPG, TCGA-ACC, TCGA-CHOL. Direct GDC project records are available at https://gdc.cancer.gov/. Single-cell RNA-seq data: GSE125449 (GEO, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125449), GSE149614 (GEO, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE149614). Spatial transcriptomics (10x Visium HCC samples): deposited under GEO accession GSEXXXXX (upon publication). miRNA target prediction databases: miRDB (https://mirdb.org/, accessed 1 Jan 2026), miRWalk (https://mirwalk.umm.de/, accessed 1 Jan 2026), DIANA-microT (https://genie.weizmann.ac.il/pubs/mirtarbase/, accessed 1 Jan 2026), ENCORI (http://rnasupport.uci.edu/databases/enori/home.html, accessed 1 Jan 2026). Drug screening: Connectivity Map (CMap) / DrugReflector (https://clue.io/, accessed 1 Jan 2026; query: BRD-K97168035, BRD-K31999759). Pathway/gene set databases: MSigDB v2024.1 (https://www.gsea-

msigdb.org/gsea/msigdb/index.jsp), KEGG (https://www.genome.jp/kegg/, accessed 1 Jan 2026), GO (Gene Ontology, https://geneontology.org/, accessed 1 Jan 2026). Protein-protein interaction network: STRING v12 (https://string-db.org/, protein accession PCSK9, combined score >= 0.7). Cancer genomics browser: cBioPortal (https://www.cbioportal.org/, datasets: TCGA-LIHC, TCGA-COAD, TCGA-READ, TCGA-STAD, TCGA-HNSC). Protein structure: PDB entry 2PMW (https://www.rcsb.org/structure/2PMW). Clinical trials: ClinicalTrials.gov (https://clinicaltrials.gov/, accessed 2 June 2026; search term: PCSK9). Code and analysis outputs supporting this manuscript are available in the project repository at the authors' institution upon reasonable request.

Ethics statement: This study used de-identified public omics and trial registry data only; no new human subjects research was conducted. Institutional review board approval was not required.

Not applicable. This study used de-identified public datasets only; no individual patient data requiring publication consent was analyzed.

Consent to publish

Not applicable. This study used de-identified public datasets; no individual patient data requiring publication consent was analyzed.

Consent to participate

Not applicable. All data used in this study were derived from publicly available de-identified datasets; no new human subjects were recruited or participated in this research.

References

1. Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. N Engl J Med. 2017;376(18):1713-1722.
2. Schwartz GG, Steg PG, Szarek M, et al. Alirocumab and cardiovascular outcomes after acute coronary syndrome. N Engl J Med. 2018;379(22):2097-2107.
3. Wang L, Li S, Luo H, et al. PCSK9 promotes the progression and metastasis of colon cancer cells through regulation of EMT and PI3K/AKT signaling in tumor cells and phenotypic polarization of macrophages. J Exp Clin Cancer Res. 2022;41(1):303.
4. Gao Y, et al. PCSK9 in liver cancers at the crossroads between lipid metabolism and immunity. Cells. 2022;11(3):432.
5. Sun C, Zhu G, Shen C, et al. Identification and validation of PCSK9 as a prognostic and immune-related influencing factor in tumorigenesis: a pan-cancer analysis. Front Oncol. 2023;13:1134063.
6. Targeting PCSK9: a promising adjuvant strategy in cancer immunotherapy. Signal Transduct Target Ther. 2021;6:296.

7. Liu X, et al. PCSK9 inhibition potentiates cancer immune checkpoint therapy. Nature. 2020;588(7839):693-698.
8. Wang R, Liu H, He P, et al. Inhibition of PCSK9 enhances the antitumor effect of PD-1 inhibitor in colorectal cancer by promoting the infiltration of CD8+ T cells and the exclusion of Treg cells. Front Immunol. 2022;13:947756.
9. Ungvari Z, Menyhart O, Lehoczki A, et al. PCSK9 expression and cancer survival: a prognostic biomarker at the intersection of oncology and geroscience. Geroscience. 2025;48(1):435-449.
10. Kim N, Kim MH, Pyo J, et al. CCR8 as a therapeutic novel target: omics-integrated comprehensive analysis for systematically prioritizing indications. Biomedicines. 2023;11(11):2910.
11. Liu J, Lichtenberg T, Hoadley KA, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. Cell. 2018;173(2):400-416.e11.
12. Sun S, Ma J, Zuo T, et al. Inhibition of PCSK9: a promising enhancer for anti-PD-1/PD-L1 immunotherapy. Research (Wash D C). 2024;7:0488.
13. Li T, Fu J, Zeng Z, et al. TIMER2.0 for tumor immune system interactions. Nucleic Acids Res. 2023;51(D1):D1066-D1070.
14. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45(W1):W98-W102.
15. Szklarczyk D, Gales L, Lyon D, et al. STRING v12: protein–protein association networks. Nucleic Acids Res. 2025;53(D1):D638-D646.
16. Uhlen M, Karlsson MJ, Zhong W, et al. A pathology atlas of the human cancer transcriptome. Science. 2017;357(6352):eaan2507.
17. Jiang P, Gu S, Pan D, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24(10):1550-1558.
18. Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
19. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18(1):220.
20. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tumor-infiltrating immune and stromal cell populations using gene expression. Genome Res. 2016;26(10):1388-1399.
21. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453-457.
22. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401-404.
23. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(269):pl1.
24. Rees MG, Seashore-Ludlow B, Cheah JH, et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. Nat Chem Biol. 2016;12(2):109-116.

25. Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(D1):D955-D961.
26. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function. J Comput Chem. 2010;31(2):455-461.
27. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021;12(1):1088.
28. Ouyang JF, Kamaraj US, Cao EY, et al. Scissor: a machine learning approach for single-cell phenotype mapping. bioRxiv. 2021;2021.09.29.462197. https://doi.org/10.1101/2021.09.29.462197
29. Okino ST, Bhatt B, Rajapakse I. scTenifoldKnk and scTenifoldNet: virtual knockout and over-expression analysis using single-cell transcriptomic data. Bioinformatics. 2021;37(21):3894-3900.
30. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19(1):15.
31. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019;9(1):5233.
32. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888-1902.e21.
33. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16(12):1289-1296.
34. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature. 2019;566(7745):496-502.
35. Wolf FA, Hamey FK, Plass M, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. 2019;20(1):59.
36. Lähnemann D, Köster J, Szczurek E, et al. scTour: a deep learning-based model for single-cell trajectory inference. bioRxiv. 2022. https://doi.org/10.1101/2022.01.24.477538
37. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv. 2018;1802.03426. https://doi.org/10.48550/arXiv.1802.03426
38. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1-2:19-25.

Tables

Tables 1 and 2 are available in the supplementary files section

Figures

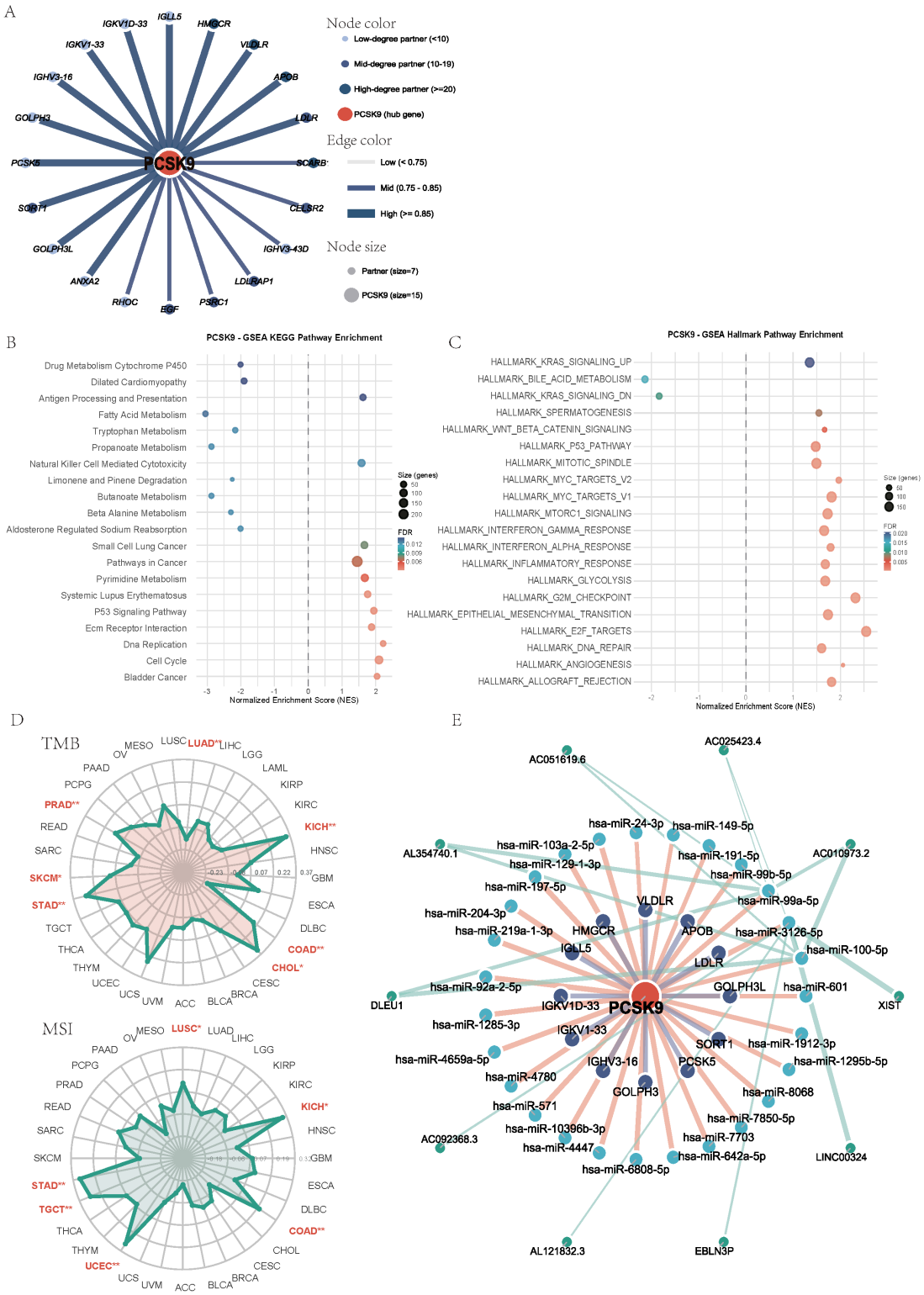


Figure 1

PCSK9 protein-protein interaction network, functional enrichment, and integrated ceRNA regulatory network. (a) PCSK9-centered PPI network from STRING v12 (combined score >= 0.7) showing PCSK9 (brick red central node) and top 20 direct interactors, with partner nodes colored by network degree (dark blue: degree >= 20; mid blue: degree >= 10; light blue: degree < 10) and edge thickness reflecting interaction confidence. (b) GSEA KEGG pathway analysis dot plot showing normalized enrichment

scores (NES) for the top 20 pathways, with color indicating FDR and size indicating gene set size. PCSK9 is positively associated with Cell Cycle (NES = 2.10, FDR = 0.003) and Bladder Cancer (NES = 2.04, FDR = 0.003) and negatively associated with Beta-Alanine (NES = -2.29) and Butanoate (NES = -2.87) Metabolism. (c) GSEA Hallmark pathway analysis showing positive enrichment in E2F Targets (NES = 2.54, FDR < 0.001), G2M Checkpoint (NES = 2.32, FDR < 0.001), MYC Targets (NES = 1.81-1.96, FDR < 0.001), and Interferon Alpha/Gamma Response (NES = 1.66-1.80, FDR < 0.001), with negative enrichment in Bile Acid Metabolism (NES = -2.13) and Fatty Acid Metabolism (NES = -1.75). (d) Prediction score heatmap of the top 20 miRNAs targeting PCSK9, integrated from miRDB, miRWalk, and ENCORI databases; color intensity represents normalized prediction confidence. (e) Integrated concentric ceRNA network of PCSK9: center = PCSK9; inner ring = PPI partners; middle ring = miRNAs predicted to target PCSK9; outer ring = lncRNAs predicted to sponge these miRNAs

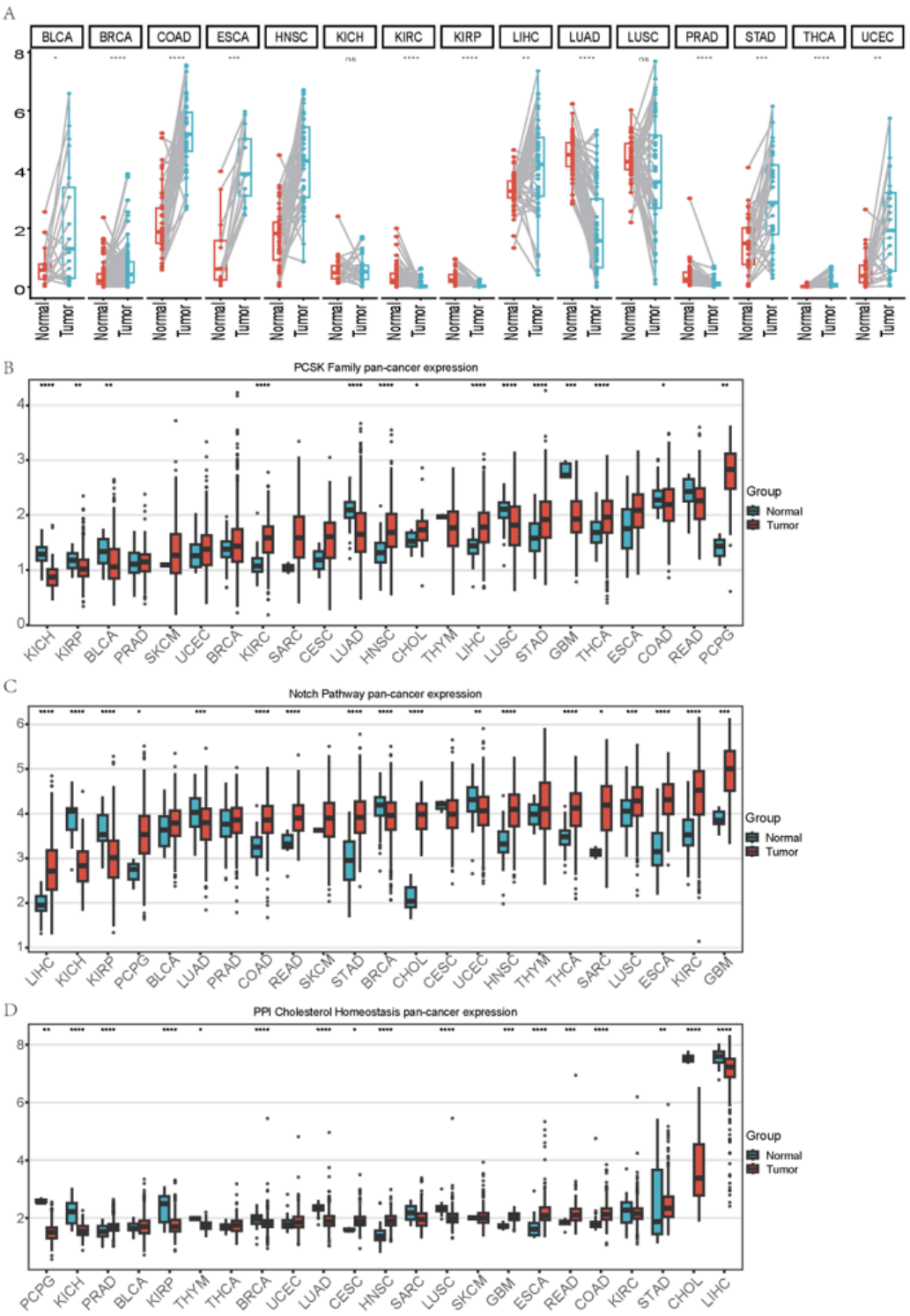


Figure 2

Pan-cancer expression profiles of PCSK9 and its functionally related pathways across TCGA. (a) PCSK9 paired tumor-normal expression boxplots across 16 TCGA cancer types (blue = normal, red = tumor, connecting lines = paired samples). PCSK9 is significantly up-regulated in COAD (log2FC = 2.71, FDR = 4.39 x 10^-18), READ, ESCA, HNSC, STAD, and LIHC, and down-regulated in KIRC and KIRP. (b) PCSK9 family member expression (PCSK1/2/4/5/6/7/9) in paired pan-cancer samples, showing distinct tumor

expression patterns: PCSK9 enriched in gastrointestinal tumors, PCSK5/6 broadly up-regulated, PCSK1 enriched in THCA. (c) Notch signaling pathway core gene expression (NOTCH1-4 receptors, JAG1-2/DLL1/3/4 ligands, HES1/HEY1 effectors), showing global Notch activation in gastrointestinal tumors and NOTCH3 down-regulation in renal cancers. (d) PPI cholesterol homeostasis pathway core gene expression (PCSK9, LDLR, APOB), showing complementary PCSK9 up-regulation and LDLR down-regulation in gastrointestinal tumors, with APOB specifically up-regulated in LIHC

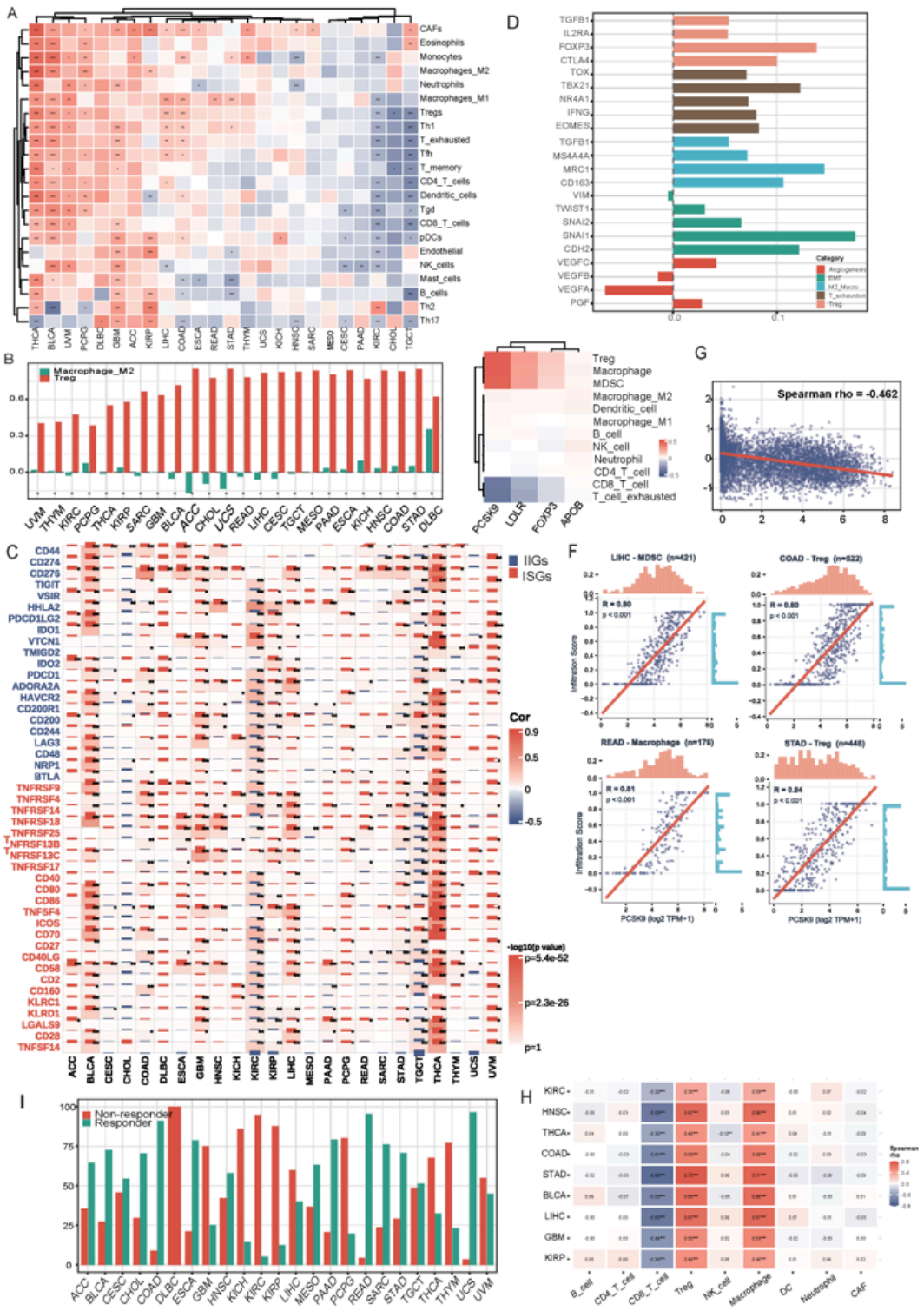


Figure 3

Pan-cancer analysis of PCSK9 expression associated with tumor immune microenvironment features and immunotherapy response. (a) Heatmap of Spearman correlations between PCSK9 expression and 22 immune cell subsets (xCell) across 24 TCGA cancer types. Red = positive, blue = negative; asterisks denote FDR significance (*P < 0.05, **P < 0.01, ***P < 0.001). PCSK9 is positively correlated with macrophages, Tregs, and MDSCs and negatively with CD8+ T cells and exhausted T cells. (b) CIBERSORT dot plot of PCSK9 correlations with Treg cells and M2 macrophages across 24 cancer types; dot size = -log10(p), color = correlation direction and magnitude. (c) Pan-cancer heatmap of PCSK9 correlations with TME functional genes grouped into T-cell exhaustion, Treg signatures, M2 macrophages, angiogenesis, and EMT categories. (d) Panorama heatmap of statistical significance (-log10(FDR)) for PCSK9-immune cell correlations across 24 cancer types. (e) Comparison of immune correlation patterns between PCSK9 and cholesterol pathway control genes (LDLR, APOB) in COAD, READ, and LIHC; LDLR shows concordant but weaker correlations, APOB shows no significant associations. (f) Representative scatter plots with marginal densities for PCSK9 vs key immunosuppressive cells in four digestive cancers: LIHC vs MDSC (R = 0.80, n = 421), COAD vs Treg (R = 0.80, n = 522), READ vs macrophages (R = 0.81, n = 176), STAD vs Treg (R = 0.84, n = 448); all P < 0.001. (g) Scatter plot of PCSK9 expression (log2(TPM+1)) vs TIDE score in the merged pan-cancer cohort (n = 6,368; Spearman rho = -0.46, P < 2.2 x 10^-16), with linear regression fit and 95% confidence interval. (h) Validation heatmap of PCSK9 correlations with nine core immune cell subsets in the independent TIDE immunotherapy cohort. (i) Comparison of predicted immunotherapy response rates between PCSK9-high and PCSK9-low groups (median-stratified) across 24 cancer types; red = higher response in PCSK9-high, blue = lower

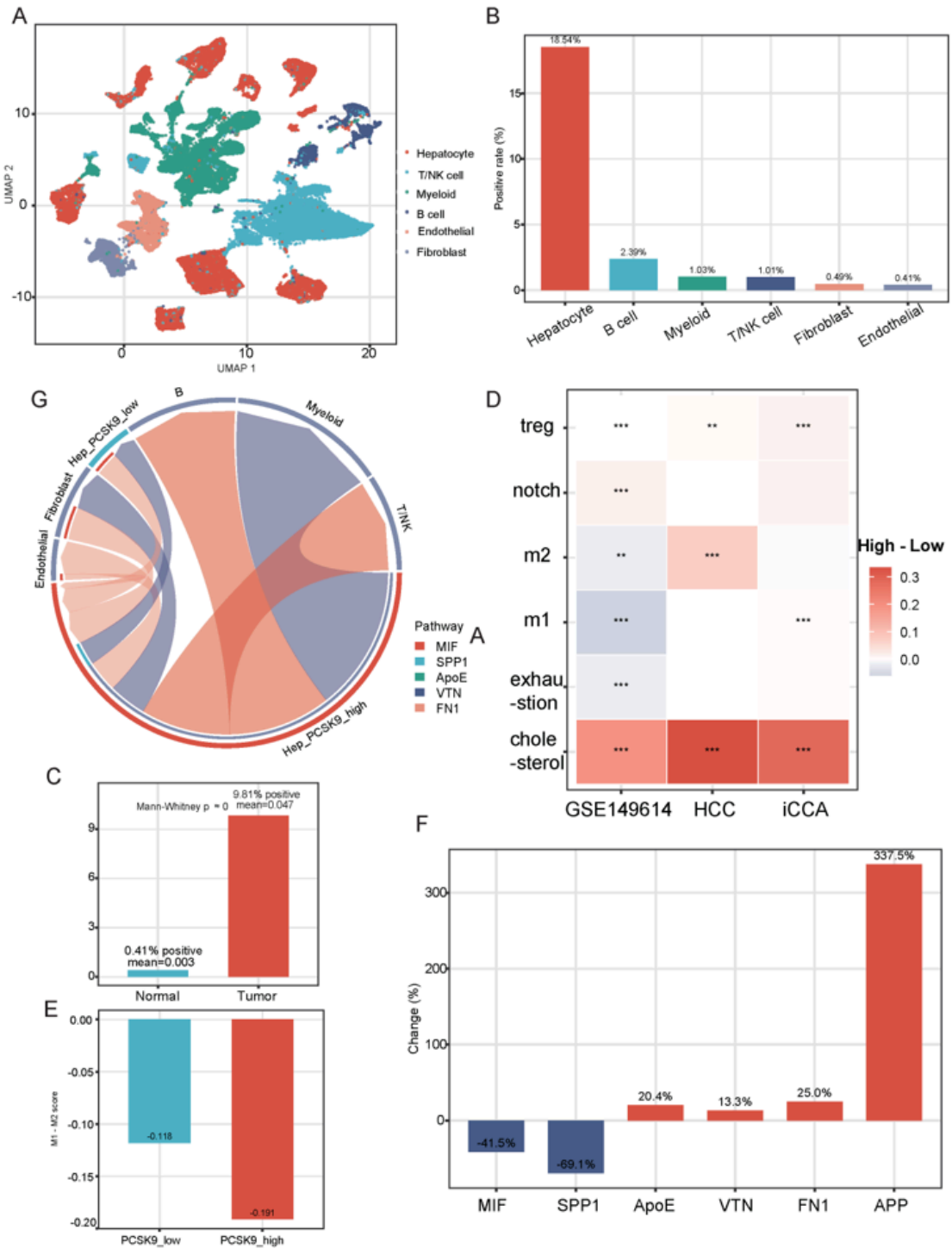


Figure 4

Single-cell landscape of PCSK9 expression and immune microenvironment shaping in hepatocellular carcinoma. (a) UMAP visualization of six major cell types in the GSE149614 dataset (71,914 cells), colored by cell type annotation: Hepatocyte (20,782), T/NK cell (25,591), Myeloid (15,947), B cell (3,685), Endothelial (3,644), Fibroblast (2,266). (b) PCSK9 expression distribution across cell types (violin+box+jitter), showing highest expression in Hepatocytes (18.54% positive, mean = 0.090) and

minimal expression in immune and stromal cells (<2.5%). (c) Comparison of PCSK9 expression between tumor and adjacent normal tissues in GSE149614, showing significant up-regulation in tumors (9.81% vs 0.41% positive, Mann-Whitney P < 2.2 x 10^-16). (d) Heatmap of immune gene set scores for PCSK9-high vs low/no hepatocytes across HCC, iCCA, and GSE149614 cohorts, including cholesterol metabolism, M1, M2, Treg, MHCII, Cytotoxic, Exhaustion, Immune_checkpoint, and NOTCH_pathway. (e) Bar plot of M1-M2 score difference in HCC hepatocytes, showing markedly lower M1-M2 in PCSK9-high cells (-0.191 vs -0.118), indicating M2-dominant phenotype. (f) Bar plot of outgoing communication differences between PCSK9-high and PCSK9-low hepatocytes in GSE149614 (CellChat), showing percentage changes for MIF, SPP1, ApoE, VTN, FN1, APP, and other key pathways. (g) CellChat communication chord diagram (PCSK9-high vs low), centered on Hep_PCSK9_high, with chord width = communication probability and color = pathway (MIF, SPP1, ApoE, VTN, FN1), illustrating reshaped communication between PCSK9-high hepatocytes and immune cells

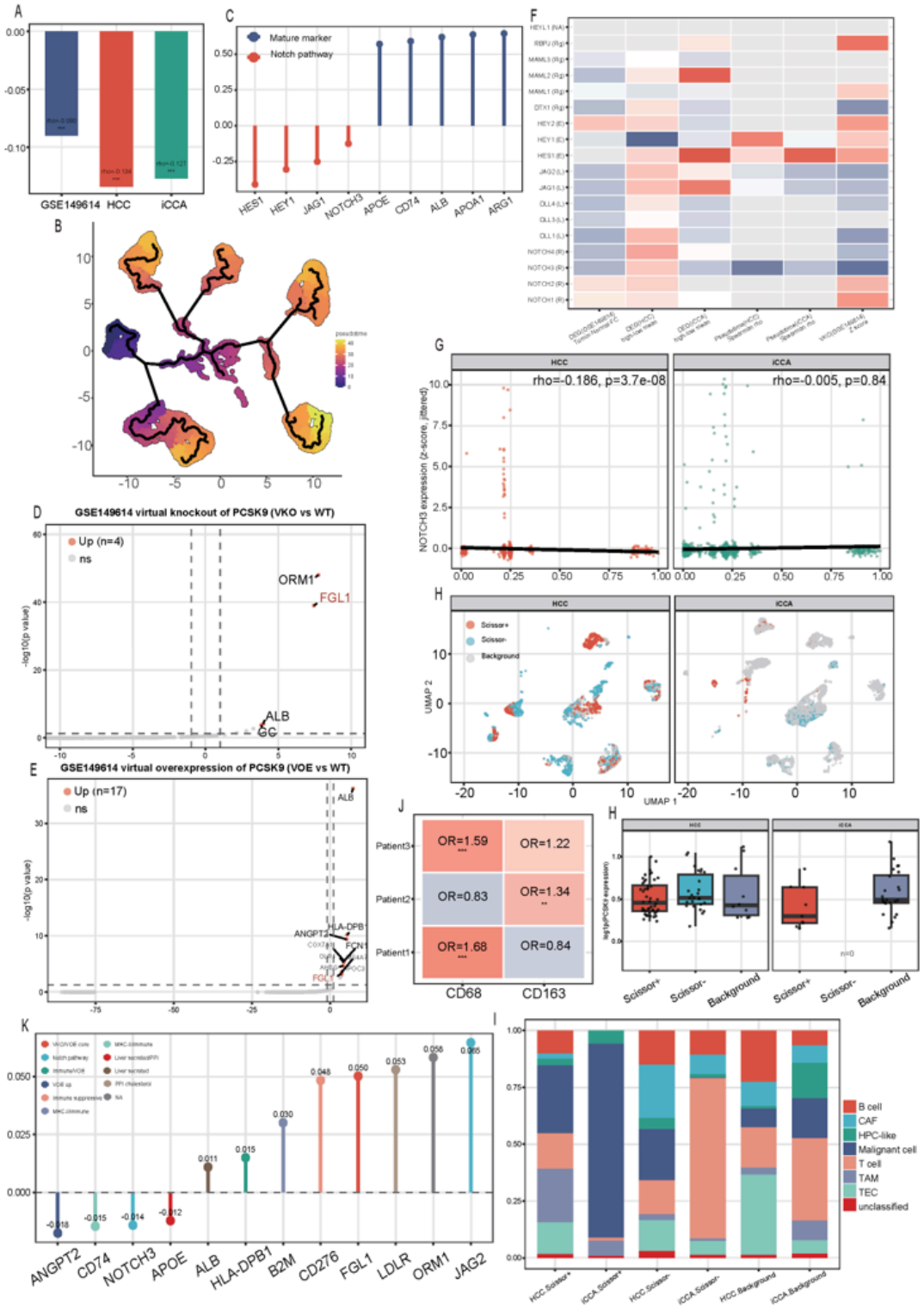
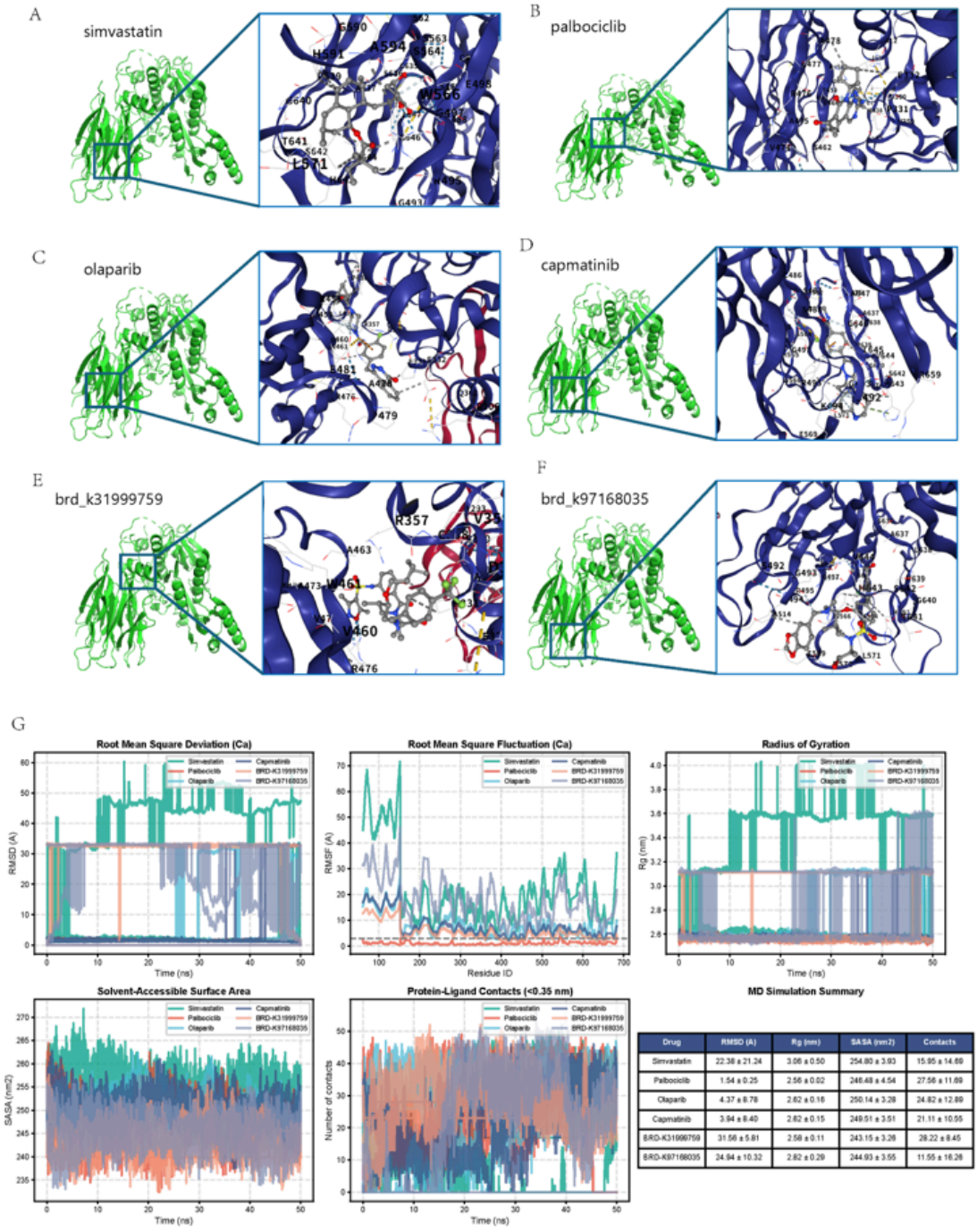


Figure 5

PCSK9-associated hepatocyte de-differentiation, in silico perturbation networks, dose-dependent Notch pathway activation, and spatial transcriptomics validation. (a) Spearman correlation between PCSK9 expression and PAGA pseudotime across three datasets (GSE149614: rho = -0.090, P = 8.88 x 10^-39; HCC: rho = -0.134, P = 7.25 x 10^-5; iCCA: rho = -0.127, P = 3.50 x 10^-8), showing consistent negative correlations. (b) Monocle 3 pseudotime UMAP of GSE149614 hepatocytes, colored by pseudotime,

rooted at the highest PCSK9-expressing cell, showing the learned principal graph trajectory. (c) Lollipop plot of Spearman correlations between pseudotime and mature hepatocyte markers (ALB, ARG1, APOA1, APOE, CD74) or Notch pathway genes (HES1, HEY1, JAG1, NOTCH3) in GSE149614 hepatocytes. (d) Volcano plot of scTenifoldKnk virtual knock-out (VKO) of PCSK9 in GSE149614, excluding PCSK9 self-perturbation, labeling significant genes including ORM1 (FC = 215.36, padj = 2.33 x 10^-46), FGL1 (FC = 173.98, padj = 1.67 x 10^-37), ALB, and GC. (e) Volcano plot of scTenifoldNet virtual over-expression (VOE) of PCSK9 in GSE149614, labeling significant genes including ALB, HLA-DPB1 (FC = 42.57, padj = 1.71 x 10^-8), ANGPT2, and FGL1. (f) Summary heatmap of Notch pathway analysis across four dimensions: DEG differences, subgroup score differences, pseudotime correlation, and VKO/VOE Z-scores; rows = Notch receptors, ligands, effectors, and regulators. (g) Scatter plot of NOTCH3 expression vs PAGA pseudotime in HCC and iCCA hepatocytes (faceted), with jitter for sparse expression; Spearman rho and P annotated (HCC: rho = -0.186, P = 3.7 x 10^-8), suggesting NOTCH3 co-localization with de-differentiation. (h) Scissor integration of TCGA-LIHC phenotypes mapped onto single-cell UMAP: HCC ImmuneScore Scissor+ cells (927, red) and iCCA OS Scissor+ cells (237, red); Scissor- and background cells shown in light cyan and gray. (i) Box plot of PCSK9 expression by Scissor group (Scissor+/Scissor-/Background) in HCC and iCCA, showing only non-zero expressing cells; Scissor+ cells exhibit significantly higher PCSK9 (HCC: P = 2.78 x 10^-9; iCCA: P = 4.33 x 10^-10). (j) Stacked bar plot of Scissor+/Scissor-/Background cell proportions across cell types in HCC and iCCA, showing HCC Scissor+ enrichment in TAMs and iCCA Scissor+ enrichment in malignant cells. (k) Heatmap of spatial co-localization between PCSK9-positive spots and macrophage markers (CD68, CD163) across three HCC Visium samples, colored by log2(odds ratio) with OR and significance annotated (*P < 0.05, **P < 0.01, ***P < 0.001). (l) Lollipop plot of mean Spearman correlation between PCSK9 and key single-cell molecules in spatial transcriptomics (three HCC Visium samples), with genes colored by functional category, showing spatial co-expression of FGL1 (rho = 0.050), JAG2 (rho = 0.065), ORM1 (rho = 0.058), and CD276/B7-H3 (rho = 0.048)



Drug	RMSD (Å)	Rg (nm)	SASA (nm2)	Contacts
Simvastatin	22.38 ± 21.24	3.06 ± 0.50	254.80 ± 3.93	15.95 ± 14.69
Palbociclib	1.54 ± 0.25	2.56 ± 0.02	246.48 ± 4.54	27.56 ± 11.89
Olaparib	4.37 ± 8.78	2.62 ± 0.16	250.14 ± 3.28	24.82 ± 12.89
Capmatinib	3.94 ± 8.40	2.62 ± 0.15	249.51 ± 3.51	21.11 ± 10.55
BRD-K31999759	31.56 ± 5.81	2.58 ± 0.11	243.15 ± 3.26	28.22 ± 8.45
BRD-K97168035	24.94 ± 10.32	2.82 ± 0.29	244.93 ± 3.55	11.55 ± 16.26

Figure 6

Drug screening, molecular docking, and molecular dynamics simulation validation of PCSK9-targeting compounds. (a) Heatmap of the top 20 compounds from DrugReflector screening of the CMap database (2,372 compounds) for reversal of the PCSK9-high gene expression signature. X-axis = compound identifiers, Y-axis = logit scores, color intensity = reversal probability. BRD-K97168035 (logit = 9.62, prob = 0.047) and BRD-K31999759 (logit = 9.21, prob = 0.031) rank first and second. (B-G) Molecular docking

binding poses of six candidate compounds with the PCSK9 catalytic domain (PDB: 2PMW): (b) simvastatin (Vina = -6.39 kcal/mol), (c) palbociclib (Vina = -7.36 kcal/mol), (d) olaparib (Vina = -8.69 kcal/mol), (e) capmatinib (Vina = -8.01 kcal/mol), (f) BRD-K31999759, (g) BRD-K97168035. Compounds shown as ball-and-stick models, PCSK9 as cartoon, key interacting residues labeled. (h) Six-complex MD simulation results (100 ns, 6 panels: RMSD, RMSF, Rg, SASA, protein-ligand close contacts, summary table). Palbociclib shows the best scaffold stability (RMSD = 1.54 ± 0.25 A, Rg = 2.56 ± 0.02 nm) with persistent close contacts (27.6 ± 11.7). (i) Four-complex MD simulation results (excluding BRD compounds, 6 panels) for focused comparison of literature-reported compounds

Supplementary Files

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

- SupplementaryFig.docx
- GraphicalAbstract.docx
- SupplementaryTables.docx
- Tables.docx