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.

ICOS Defines an Early Regulatory Exhaustion Immune Axis in HBV-Associated Hepatocellular Carcinoma

Grifton Tafadzwa Muchovu

20063791@studenti.uniupo.it

University of Eastern Piemonte https://orcid.org/0009-0001-4465-4116

Concorde Niyigaba Isange

University of Eastern Piemonte https://orcid.org/0009-0008-3528-2980

Jean Francois Regis Igiramaboko2

https://orcid.org/0009-0001-6096-7622

Cephas Mawere

Harare Institute of Technology https://orcid.org/0009-0009-4833-1929

Research Article

Keywords: ICOS, Hepatocellular carcinoma, Single-cell RNA sequencing, T-cell differentiation, Immune exhaustion, Tumor microenvironment

Posted Date: February 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8971146/v1

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

Additional Declarations: The authors declare no competing interests.

Abstract

HBV-associated hepatocellular carcinoma (HCC) develops within a chronically inflamed tumor microenvironment. It is defined by persistent T-cell activation and checkpoint remodeling. The biological role of inducible T-cell co-stimulator (ICOS) in this context is still incompletely defined. The aim of this study was to evaluate whether ICOS defines a distinct T-cell transcriptional state, its position in T-cell differentiation, and its tumor-level immune relevance in HBV-related HCC. We performed a two-phase integrative transcriptomic analysis combining single-cell RNA sequencing of tumor-infiltrating T cells (GSE149614; 20,633 cells) with independent validation in TCGA-LIHC bulk RNA-seq data (n = 729 tumors). ICOS-based stratification was followed by module scoring, differential expression analysis, Hallmark pathway enrichment, diffusion pseudotime inference, multivariable regression, mediation modeling, and survival analysis. ICOS expression was confined to a minority subset (10%) and non-randomly enriched in discrete CD4-dominant clusters (up to 39.8%). $ICOS^{high}$ cells exhibited amplified regulatory transcriptional activity (Treg_score 0.375 vs − 0.041; ρ = 0.165; $p = 1.06 \times 10^{-125}$) with more moderate exhaustion enrichment. On the other hand, CD8-restricted analysis showed no significant difference in exhaustion (p = 0.384). Differential expression revealed extensive transcriptional reprogramming (98.4% of tested genes at FDR < 0.05). Hallmark analysis exhibited coordinated enrichment of 30/50 pathways in $ICOS^{low}$ cells, showing relative transcriptional specialization rather than global activation in $ICOS^{high}$ cells. Diffusion pseudotime positioned $ICOS^{high}$ cells in early differentiation states (61.5% lower median pseudotime; $p = 6.47 \times 10^{-52}$), with exhaustion intensity declining along trajectory (ρ = −0.3601; $p < 10^{-300}$). At the tumor scale, ICOS strongly correlated with exhaustion (ρ = 0.727; $p = 8.52 \times 10^{-87}$) and remained the dominant independent predictor (β = 0.710; $R^2 = 0.542$), with partial CD8 mediation (33.7%). ICOS did not independently predict overall survival, and it defines an early-positioned, transcriptionally coherent, regulatory-biased immune state that coordinates activation–regulation–exhaustion dynamics in HBV-associated HCC rather than marking terminal dysfunction.

Introduction

Hepatocellular carcinoma (HCC) is still a leading cause of cancer deaths worldwide [1]. In many areas, long-term infection with Hepatitis B virus is the main cause [2, 3]. This leads to a tumor microenvironment that is chronically inflamed and complex [4, 5]. The environment is marked by ongoing T-cell activation, gradual changes in function, and shifts in immune regulation [4, 6]. Immune checkpoint therapies have improved outcomes for some patients with advanced HCC [7, 8], but responses vary, highlighting gaps in our understanding of T-cell states and regulation within tumors [5, 7].

Tumor-infiltrating T cells in HCC usually coexpress activation and inhibitory receptors [5, 9]. Emerging evidence suggests that exhaustion does not signify a binary end-stage dysfunction. Rather, it shows a structured, plastic differentiation spectrum driven by chronic antigen stimulation [9, 10]. Within this spectrum, regulatory T cells (Tregs) and checkpoint-expressing subsets have important roles in directing immune tolerance [5, 11]. However, the core transcriptional drivers that coordinate activation, regulation, and exhaustion programs within HBV-associated HCC remain incompletely understood [6, 12].

Inducible T-cell co-stimulator (ICOS) is a costimulatory receptor that belongs to the CD28 family [13]. It has been identified as an important modulator of T-cell activation, survival, and differentiation [14]. ICOS signaling activates the PI3K–AKT pathway and enhances follicular helper T-cell differentiation. Concurrently, it also exerts immunosuppressive effects [14, 15]. In other cancer contexts, ICOS expression has also been reported in both effector and regulatory T-cell compartments [16, 17].

Previous work on ICOS in cancer has depended on bulk transcriptomic analysis or limited phenotypic profiling [16, 18]. This masks the precise maturation stage and transcriptional heterogeneity of ICOS-expressing cells. The introduction of single-cell RNA sequencing (scRNA-seq) has now allowed the mapping of high-resolution immune cell states and a spectrum of cellular development within tumors [5, 19]. Furthermore, computational tools such as trajectory inference allow us to reconstruct the developmental trajectories of immune cells [20]. This offers a mechanistic insight into how costimulatory signaling interfaces with exhaustion and regulatory programming.

ICOS is mechanistically associated with immune regulation. However, its transcriptional architecture, differentiation dynamics, and tumor-level functional coupling in HBV-related HCC have not been systematically integrated into a organized framework [6, 16]. Precisely, it is unclear whether ICOS defines a stable suppressive lineage or a transient activation state. It is also still unclear whether it is a central regulatory hub coordinating the balance between activation and exhaustion within the tumor microenvironment.

We hypothesized that ICOS expression identifies a specific transcriptionally distinct, T-cell state enriched for regulatory programs. We suspect that it occupies a defined position along the tumor-associated differentiation continuum. We also hypothesized that this ICOS-associated immune state would show measurable coupling with exhaustion intensity at the tumor scale and would be translationally relevant when examined in independent patient cohorts.

To test this hypothesis, we performed an integrative two-phase analysis. First, we analyzed single-cell RNA sequencing data from HBV-associated HCC tumors (GSE149614). We performed ICOS-based stratification, differential expression modeling, and Hallmark pathway enrichment analysis. We also inferred diffusion-based pseudotime trajectories. In the second phase, we validated the ICOS-associated immune axis in the TCGA-LIHC cohort using bulk RNA-seq–derived immune scoring, multivariable modeling, mediation analysis, interaction testing, and survival analysis.

By integrating these techniques, this study aims to clarify the biological significance of ICOS expression in HCC. It also seeks to position ICOS within a structured immune regulatory framework relevant to chronic viral malignancy and immunotherapeutic targeting.

Methodology

Study Design Overview

This study was conducted as a two-phase integrative transcriptomic investigation. The first phase comprised a discovery analysis using single-cell RNA sequencing (scRNA-seq) of tumor-infiltrating lymphocytes derived from hepatitis B virus (HBV)–associated hepatocellular carcinoma (HCC). This phase included strict quality control, lineage restriction, dimensionality minimization, unsupervised clustering, ICOS-based stratification, lineage-specific functional scoring, differential gene expression modeling, pathway enrichment analysis, and trajectory inference. The second phase consisted of independent validation utilizing bulk RNA sequencing data from the TCGA-LIHC cohort to evaluate the clinical and survival relevance of the ICOS-associated immune phenotype beyond the viral context.

The chief aim was to determine whether ICOS expression defines a transcriptionally coherent regulatory–exhausted T-cell state within the tumor microenvironment and whether this immune axis exhibits translational significance at the patient level.

All analyses were conducted in Google Colab using Python (v3.12). Core libraries included Scanpy (single-cell analysis), AnnData (structured data storage), Pandas and NumPy (data handling), SciPy (nonparametric testing), Statsmodels (regression modeling), Scikit-learn (feature scaling), and Matplotlib/Seaborn (visualization). All statistical tests were two-sided. Nominal significance was defined at $\alpha = 0.05$. Multiple hypothesis testing was controlled using the Benjamini–Hochberg false discovery rate (FDR) procedure where applicable.

Data Acquisition and Preprocessing (Single-Cell RNA Sequencing)

Raw scRNA-seq count matrices and metadata were obtained from the Gene Expression Omnibus (GEO) under accession GSE149614. The dataset included 71,915 tumor-derived cells profiled using droplet-based single-cell profiling. Metadata contained barcode identifiers, sample IDs, patient identifiers, tumor stage, viral status, and pre-annotated cell type labels.

To focus specifically on adaptive immune programs, cells annotated as “T/NK” were extracted. Selective barcode parsing was used to maintain computational effectiveness. The resulting AnnData object contained 25,590 cells and 25,712 genes.

Quality control was performed using Scanpy. Cells with fewer than 200 detected genes, elevated mitochondrial transcript content, or low transcriptomic complexity were excluded. Genes expressed in fewer than three cells were removed. Iterative clustering followed by marker-based inspection identified clusters enriched for myeloid markers (LYZ, LST1, AIF1, HLA-DPA1), which were excluded to eliminate lineage contamination. After filtration, 20,633 high-confidence tumor-infiltrating T cells were retained.

Although the matrices from GEO were already normalised, standardized preprocessing was re-applied to guarantee consistency. Library-size normalization to 10,000 counts per cell was followed by log1p transformation. Highly variable genes (HVGs) were selected using dispersion-based criteria (min_mean = 0.0125; max_mean = 3; min_disp = 0.5). This resulted in 3,926 HVGs. Data were scaled to unit variance, with a maximum value of 10, to reduce outlier influence.

Principal component analysis (PCA) was performed on HVGs, and the top 30 principal components were retained based on variance stabilization and elbow plot inspection. A k-nearest neighbor graph (n_neighbors = 15; n_pcs = 30) was constructed, followed by Leiden clustering (resolution = 0.6; two iterations). UMAP embedding was generated for two-dimensional visualization.

Lineage Resolution and CD4/CD8 Stratification

To resolve major T-cell lineages, canonical markers CD4, CD8A, and CD8B were measured at both the single-cell and cluster levels. Mean normalized expression per cluster was calculated. Clusters were classified as CD4-dominant if mean CD4 expression exceeded CD8A, and CD8-dominant otherwise. This yielded 11,316 $CD4^+$ T cells (54.9%) and 9,317 $CD8^+$ T cells (45.1%).

Given the established role of $CD8^+$ T cells in cytotoxic and exhaustion-related programs, lineage-restricted analyses were performed. CD8-specific subsets were extracted, and ICOS-associated transcriptional comparisons were repeated within this restricted compartment to determine whether ICOS–exhaustion coupling was CD8-dependent.

ICOS Quantification and Stratification

ICOS expression was quantified per cell from the normalized matrix. Due to the typical zero inflation in scRNA-seq data, percentile thresholds below the 90th percentile collapsed to background levels. We defined 'ICOS-high' cells as the 90th percentile of ICOS expression across the entire T-cell compartment. Cells that satisfied this threshold were labeled ICOS-high; all others were ICOS-low.

To guarantee robustness, other stratifications were also performed. These include binary ICOS positivity (> 0 vs 0 expression) and a stricter ≥ 95th percentile definition. Across all definitions, enrichment of regulatory and exhaustion programs remained statistically significant, confirming threshold stability.

Patient-Level Distribution Validation

To exclude single-patient dominance effects, ICOS-high proportions were calculated for each tumor sample. Metadata were aligned at the barcode level, and per-patient proportions were calculated. ICOS-high cells were seen across all independent tumors (range 1.7%–12.6%), demonstrating that ICOS enrichment was not attributable to isolated patient-level artifacts.

Transcriptional Signature Scoring

Regulatory T-cell (Treg) programs were quantified using FOXP3, IL2RA, CTLA4, IKZF2, and TNFRSF18. Exhaustion signatures included PDCD1, LAG3, TIGIT, HAVCR2, TOX, and ENTPD1. Module scores were computed using Scanpy’s score_genes function.

Group comparisons between ICOS-high and ICOS-low cells were performed using the Wilcoxon rank-sum test. Effect size was reported as median difference (High − Low). FDR correction was applied.

CD8-restricted analyses were conducted separately to determine lineage-specific associations.

Differential Gene Expression and Pathway Enrichment

Differential expression between ICOS-high and ICOS-low cells was performed using Wilcoxon rank-sum testing with FDR correction. Genes were considered significant if adjusted p-value < 0.05 and |log2FC| > 0.25.

Pre-ranked gene lists were subjected to Gene Set Enrichment Analysis (GSEA) using the MSigDB Hallmark collection (1,000 permutations; gene set size 15–500). Consistent with standard guidelines for exploratory transcriptomic analyses, FDR q < 0.25 was considered significant.

Trajectory Inference

Diffusion-based pseudotime (DPT) analysis was performed to evaluate differentiation positioning. PCA and kNN graphs were recomputed prior to diffusion mapping to prevent topology bias. Root cells were defined within early activation clusters based on low exhaustion scores.

Median pseudotime differences between ICOS groups were assessed using the Mann–Whitney U test. Spearman's correlation was used to evaluate associations between pseudotime and exhaustion signature scores. Pseudotime was binned into deciles to evaluate non-linear dynamics.

Bulk RNA-Seq Validation (TCGA-LIHC)

To validate the ICOS-associated immune phenotype at the tumor level, bulk RNA sequencing data from the TCGA Liver Hepatocellular Carcinoma (TCGA-LIHC) PanCancer Atlas cohort were retrieved via cBioPortal. Gene expression values were kept on the normalized RSEM scale without subsequent adjustments to ensure fair comparability across samples. Sample identifiers were harmonized to patient-level TCGA barcodes (first 12 characters), and expression data were merged with clinical annotations using inner joins to ensure complete case matching.

To mechanistically position ICOS within the tumor immune microenvironment, composite immune functional scores were constructed. A T cell exhaustion score was calculated for each tumor as the mean normalized expression of canonical inhibitory receptors (PDCD1, CTLA4, LAG3, TIGIT, and HAVCR2). A cytotoxic score was similarly computed as the mean expression of CD8A, GZMB, and PRF1 to approximate effector T cell activity. These composite scores were calculated using vectorized operations in Python.

Spearman's rank correlation analysis was executed to evaluate monotonic associations between ICOS expression and immune functional markers. The markers include exhaustion score, CD8A, PDCD1, CTLA4, and FOXP3. To evaluate dose-dependent relationships, tumors were stratified into quartiles based on ICOS expression using quantile-based binning (pandas.qcut), and global differences in exhaustion score across quartiles were assessed using the non-parametric Kruskal–Wallis test.

To determine whether ICOS independently predicts exhaustion intensity beyond cytotoxic infiltration, multivariable ordinary least squares (OLS) regression modeling was performed with the following specification: Exhaustion Score ~ ICOS + Cytotoxic Score. Predictors were standardized using z-score transformation (StandardScaler, scikit-learn). This allows direct comparison of effect sizes. We used (R^2), standardized regression coefficients (β), and p-values to measure our results. We also ran regression diagnostics to evaluate residual normality and multicollinearity.

To formally assess pathway structure, a mediation analysis was conducted to test whether CD8A expression partially mediates the ICOS–exhaustion relationship. Linear models were fitted for ICOS → CD8A and Exhaustion ~ ICOS + CD8A. Indirect (a × b) and direct (c') effects were estimated, and bootstrapped confidence intervals were generated using 5,000 resamples to assess stability.

An interaction model (Exhaustion ~ ICOS + CD8A + ICOS × CD8A) was also used to measure synergistic amplification of exhaustion levels.

For clinical validation, overall survival (OS) duration in months and event status were obtained from clinical annotations. Samples lacking complete OS data were excluded (final n = 371). Cox proportional hazards regression was performed using continuous ICOS and FOXP3 expression. Hazard ratios (HRs) with 95% confidence intervals were reported. Kaplan–Meier survival curves were generated using median-based dichotomization and compared via the log-rank test.

All tests followed a standard 5% significance level. Since we were observing multifaceted data, we've interpreted our results as related trends rather than assuming a direct chain of cause and effect.

Results

ICOS Expression Is Heterogeneous and Non-Randomly Distributed Across the Tumor-Infiltrating T-Cell Landscape

Unsupervised Leiden clustering resolved the tumor-infiltrating T-cell compartment into 27 transcriptionally distinct clusters. This reveals a highly structured immune landscape (Fig. 1A). Overlay of ICOS transcript abundance onto the UMAP embedding demonstrated spatially restricted expression rather than diffuse stochastic distribution (Fig. 1C). ICOS expression is particularly distributed among distinct transcriptional compartments. This indicates a state-specific enrichment.

Stratification using a 90th percentile expression threshold defined 2,064 $ICOS^{high}$cells (10.0%) and 18,569 $ICOS^{low}$ cells (90.0%), confirming that ICOS expression is confined to a minority T-cell subpopulation. Importantly, $ICOS^{high}$ cells were non-uniformly distributed across clusters. Cluster-level quantification revealed marked enrichment in specific transcriptional states (Fig. 1D; Supplementary Table S1). Cluster 10 exhibited the highest $ICOS^{high}$ proportion (39.8%), followed by clusters 11 (24.7%) and 9 (21.3%), whereas the majority of clusters demonstrated $ICOS^{high}$ frequencies below 5%. Several clusters, including 14, 15, 22, 23, 25, and 26, contained fewer than 2% $ICOS^{high}$ cells, confirming highly non-random localization.

Lineage mapping further demonstrated preferential enrichment of $ICOS^{high}$ cells within the $CD4^+$ compartment (Fig. 1B; Supplementary Table S2). CD4-dominant regions displayed significantly higher ICOS signal intensity than CD8-restricted clusters, supporting a regulatory-biased spatial distribution. Collectively, these data indicate that ICOS marks a spatially and transcriptionally restricted T-cell state rather than a uniform activation across the compartment.

$ICOS^{high}$ *Cells Exhibit Amplified Regulatory and Checkpoint-Associated Programs*

To determine whether ICOS enrichment reflects functional specialization, we performed module scoring using a canonical regulatory T-cell gene set (FOXP3, IL2RA, CTLA4, IKZF2, TNFRSF18). $ICOS^{high}$ cells showed markedly higher regulatory activity compared with $ICOS^{low}$ cells (Fig. 2A–B). Mean Treg_score was 0.375 in $ICOS^{high}$ cells versus − 0.041 in $ICOS^{low}$ cells, representing an approximate 9.1-fold relative shift in regulatory transcriptional activity. Statistical separation was highly significant (Wilcoxon test, FDR < 0.05).

Continuous modeling confirmed this association. Spearman correlation between ICOS expression and Treg_score yielded $\rho = 0.165$ ($p = 1.06 \times 10^{-125}$). This demonstrates strong monotonic coupling between ICOS abundance and regulatory transcriptional programming.

Cluster-level integration reinforced this relationship. Cluster 10—the most ICOS-enriched cluster (39.8%)—also demonstrated the highest mean Treg_score (1.60), indicating spatial co-localization of ICOS expression and regulatory transcriptional identity. Scoring of exhaustion-associated checkpoints (PDCD1, LAG3, HAVCR2, TIGIT, TOX, ENTPD1) showed a statistically significant yet quantitatively more moderate enrichment in $ICOS^{high}$ cells. Mean exhaustion score was 0.245 in $ICOS^{high}$ cells compared to − 0.027 in $ICOS^{low}$ cells (Mann–Whitney $p = 5.22 \times 10^{-68}$), with Spearman correlation $\rho = 0.121$ ($p = 1.70 \times 10^{-68}$).

An analysis restricted to CD8 cells (n = 9,317 $CD8^+$ cells) showed no significant difference in exhaustion between $ICOS^{high}$ and $ICOS^{low}$ CD8 cells ($p = 0.384$). This shows that ICOS-associated transcriptional reprogramming is not driven only by classical CD8 exhaustion but rather reflects broader immunoregulatory programming.

Taken together, these findings position ICOS^{high} cells as transcriptionally regulatory rather than exhausted.

Differential Expression Analysis Shows Widespread Transcriptional Transformation in ICOS ^{high} Cells.

To evaluate transcriptional divergence between ICOS^{high} and ICOS^{low} populations, we performed differential gene expression analysis across 4,395 genes (Supplementary Table S3). A total of 4,324 genes (98.4%) were significantly altered at FDR < 0.05. This indicates a global transcriptional reprogramming rather than a focal perturbation. The top 50 upregulated and downregulated genes are summarized in Supplementary Table S6.

The volcano plot showed strong right-skewed enrichment in ICOS^{high} cells (Fig. 3A). As well, numerous transcripts showed high statistical significance (adjusted p-values exceeding 10^{-300}). Top upregulated transcripts included ICOS, which served as an internal validation control, S100A4 (adjusted p = 6.90×10^{-228}), LTB (1.09×10^{-85}), IL7R (1.08×10^{-18}), and ANXA1 (1.55×10^{-18}).

Canonical regulatory and checkpoint-associated genes were also significantly enriched. These include CTLA4 (3.88×10^{-21}), TIGIT (1.81×10^{-9}), PDCD1 (4.16×10^{-6}), IL2RA (1.06×10^{-4}), and FOXP3 (0.039). These data confirm internal biological coherence and validate the ICOS^{high} classification strategy.

Ranked gene heatmap visualization demonstrated clear block segregation between ICOS^{high} and ICOS^{low} populations (Supplementary Table S4), indicating coordinated transcriptional identity rather than stochastic noise. Collectively, these data demonstrate that ICOS^{high} cells represent a globally reprogrammed regulatory transcriptional state within the tumor-infiltrating T-cell compartment.

Hallmark Pathway Analysis Shows Coordinated Downregulation of Canonical Programs in ICOS^{high} Cells

Given the extensive transcriptional divergence observed between ICOS^{high} and ICOS^{low} cells, we sought to contextualize these global shifts at the pathway level. We used pre-ranked Gene Set Enrichment Analysis across the 50 Hallmark pathways. Amongst the50 pathways evaluated, 30 (60.0%) satisfied significance at FDR q < 0.05 (Supplementary Table S5). Notably, all significant pathways showed negative normalized enrichment scores (NESs ranging from − 5.06 to − 1.55). This indicates a coordinated enrichment in ICOS^{low} cells relative toICOS^{high}cells (Fig. 3B).

The strongest enrichment signals included were observed in:

- -HALLMARK COAGULATION (NES = − 5.06; FDR = 0.000)
- -HALLMARK XENOBIOTIC METABOLISM (NES = − 3.83),
- -HALLMARK EPITHELIAL_MESENCHYMAL TRANSITION (NES = − 3.27),

- -HALLMARK UV RESPONSE UP (NES ≈ − 3.1), and HALLMARK ADIPOGENESIS (NES ≈ − 3.0). Additionally, immune-relevant pathways were significantly enriched in $ICOS^{low}$ cells. These includedTNFA SIGNALING VIA NFKB, TGF BETA SIGNALING, P53 PATHWAY, and APOPTOSIS.

Importantly, none of the Hallmark pathways demonstrated a positive NES that met FDR significance. Thus, $ICOS^{high}$ cells do not activate a dominant canonical inflammatory, proliferative, or metabolic Hallmark program. Instead, ICOS expression is associated with relative suppression of broad metabolic and stress-response pathways, suggesting that $ICOS^{high}$ cells occupy a transcriptionally restrained or specialized regulatory state rather than a globally activated phenotype. For complete transparency and full ranked pathway output across all 50 Hallmark gene sets, see Supplementary Table S8.

$ICOS^{high}$ Cells Occupy Early Differentiation States Along the Diffusion Trajectory

To determine whether ICOS-associated transcriptional specialization corresponds to a defined differentiation state, we performed diffusion pseudotime analysis across the T-cell compartment. Distributional comparison revealed that $ICOS^{high}$ cells localize preferentially to early trajectory states (Fig. 4C). Median pseudotime for $ICOS^{high}$ cells was 0.01099 compared with 0.02853 for $ICOS^{low}$ cells, representing a 61.5% relative reduction (Mann–Whitney $p = 6.47 \times 10^{-52}$). This separation was not subtle: the interquartile range of $ICOS^{high}$ cells was tightly compressed toward early states, whereas $ICOS^{low}$ cells spanned the full differentiation continuum, including terminal regions. Median pseudotime comparisons and statistical summaries are provided in Supplementary Table S7.

To put context on the functional dynamics along this trajectory, we modelled exhaustion scores across pseudotime bins (Fig. 4A). We also performed a Spearman analysis, which showed a strong negative monotonic association between pseudotime and exhaustion intensity ($\rho = -0.3601$, $p < 1 \times 10^{-300}$). This indicates a progressive decline in exhaustion-associated transcriptional programs along the inferred differentiation continuum. Decile modeling revealed an approximate 170% decrease in mean exhaustion intensity from the earliest to the latest pseudotime bins. This supports a continuous rather than binary transition. These data show that terminal differentiation in this dataset is not synonymous with maximal exhaustion. It also shows a mechanistic distinction between differentiation state and checkpoint-associated transcriptional intensity.

Modeling the proportion of $ICOS^{high}$ cells across pseudotime bins demonstrated clear enrichment in early differentiation states (Fig. 4B). Early bins (0–10% pseudotime) contained approximately 14–15% $ICOS^{high}$ cells, intermediate bins showed 8–10%, and late bins declined to ~ 3%. This represents a 4–5 fold relative enrichment in early trajectory states compared with late states. Collectively, these results position ICOS expression within an early, regulatory-biased differentiation state rather than a terminally exhausted phenotype.

Bulk Tumor Transcriptomics Confirms a Coordinated ICOS–Immune Regulatory Axis

To determine whether single-cell findings apply to tumor-level biology, we interrogated TCGA-LIHC bulk RNA-seq data (n = 729 tumors). Spearman correlation analysis indicated strong monotonic coupling between ICOS expression and multiple immune functional axes (Fig. 5B; Supplementary Table S9). ICOS strongly correlated with:

-exhaustion score ($\rho = 0.727$, $p = 8.52 \times 10^{-87}$), CTLA4 ($\rho = 0.760$, $p = 3.00 \times 10^{-99}$), PDCD1 ($\rho = 0.661$, $p = 1.15 \times 10^{-66}$), CD8A ($\rho = 0.617$, $p = 6.62 \times 10^{-56}$), FOXP3 ($\rho = 0.299$, $p = 3.39 \times 10$).

ICOS shows stronger coupling to exhaustion-associated programs than to FOXP3 alone. This suggests a wider immune regulatory coordination beyond classical Treg markers.

Quartile stratification further revealed a clear dose-dependent gradient (Fig. 5A). Median exhaustion scores increased stepwise across ICOS quartiles: ~4.0 (Q1), ~ 4.8 (Q2), ~ 5.1 (Q3), and ~ 6.5 (Q4), representing an approximately 62% relative increase from lowest to highest ICOS quartile (Kruskal–Wallis $p = 8.97 \times 10^{-63}$). This monotonic gradient supports a quantitative ICOS–exhaustion axis at the tumor scale.

Multivariable regression modeling (Exhaustion ~ ICOS + Cytotoxic score) demonstrated strong explanatory power ($R^2 = 0.542$; model $p < 1 \times 10^{-124}$). Standardized beta coefficients confirmed ICOS as the dominant independent predictor ($\beta = 0.710$) compared with cytotoxic score ($\beta = 0.487$) (Fig. 5D;Supplementary Table S10). These data show that ICOS retains independent predictive capacity beyond cytotoxic infiltration.

Mediation analysis also revealed partial dependence on CD8. The indirect effect of ICOS through CD8 (ICOS → CD8 → Exhaustion) was 0.189 (95% CI: 0.132–0.244), accounting for 33.7% of the total effect, while a substantial direct ICOS effect persisted (0.372) (Supplementary Table S12). Interaction modeling confirmed synergistic amplification between ICOS and CD8 (ICOS × CD8 interaction $p = 0.007$; Supplementary Table S11), supporting a cooperative regulatory network rather than simple collinearity.

ICOS Does Not Independently Predict Overall Survival in Untreated Cohorts

Despite strong immune coupling, survival analysis in TCGA-LIHC (n = 371 evaluable patients) demonstrated no independent prognostic effect. Kaplan–Meier analysis stratified by ICOS expression showed no significant survival difference (log-rank $p = 0.6372$) (Fig. 6A). Cox regression confirmed the absence of association (ICOS HR = 0.99, $p = 0.76$) (Supplementary Table S15.

Combined ICOS–FOXP3 stratification likewise demonstrated no survival difference (log-rank $p = 0.4405$) (Fig. 6B), and FOXP3 was not independently prognostic (HR = 0.97, $p = 0.37$). Thus, neither ICOS alone nor ICOS–FOXP3 co-stratification predicts overall survival in untreated tumors.

Threshold sensitivity analysis confirmed robustness of immune coupling across alternative ICOS cutoffs (Supplementary Tables S13–S14). Exhaustion effect sizes increased with stricter thresholds (Δ median = 0.041 for ICOS > 0; 0.061 for ≥ 90%; 0.171 for ≥ 95%), whereas Treg-associated effects remained stable (~ 0.015–0.020). These results show that ICOS-driven exhaustion is most intense at peak expression levels, whereas the cell's regulatory functions remain stable regardless of the ICOS concentration.

Integrated Interpretation

Across single-cell trajectory modeling and bulk tumor transcriptomics, ICOS expression marks early differentiation positioning, regulatory-biased transcriptional programming, strong exhaustion coupling at tumor scale, and independent predictive capacity beyond cytotoxic infiltration. However, ICOS does not independently determine overall survival in untreated cohorts.

Collectively, this evidence defines ICOS not as a terminal exhaustion marker nor a standalone prognostic biomarker, but as a central immune regulatory hub coordinating activation, regulation, and exhaustion dynamics amid the tumor microenvironment.

Discussion

This study provides a layered dissection of ICOS-associated T-cell states within HBV-related hepatocellular carcinoma (HCC). It integrates single-cell transcriptional profiling, pathway enrichment, differentiation trajectory inference, and independent bulk tumor validation. In these detailed analytical layers, ICOS expression consistently denotes a transcriptionally coherent, regulatory-biased, early-positioned immune state. This immune state is highly correlated to exhaustion intensity at the tumor scale but does not independently dictate survival in untreated cohorts. These findings improve the biological interpretation of ICOS in chronic viral malignancy and challenge the simplistic equation of ICOS enrichment with terminal exhaustion [6, 16].

ICOS Defines a Spatially Restricted and Transcriptionally Coherent T-Cell State

Single-cell clustering identified 27 distinct transcriptional states within the tumor-infiltrating T-cell compartment. This shows a highly structured immune ecosystem [5, 12, 19]. Within this microenvironment, ICOS expression was confined to a minority (~ 10%) subpopulation and was non-randomly enriched in discrete clusters. It was most prominent within CD4-dominant compartments. This spatial restriction is consistent with established ICOS biology, which shows that ICOS is preferentially induced upon TCR engagement and costimulatory signaling, mainly within $CD4^+$ T-cell subsets including T follicular helper (Tfh) and regulatory T cells (Tregs) [13–15].

ICOS enrichment was not diffuse or evenly distributed, arguing against the interpretation that ICOS merely reflects generalized activation. Instead, cluster-level enrichment and lineage mapping showed that ICOS marks defined transcriptional state within the tumor microenvironment. This supports a model

in which ICOS expression is tightly regulated and context-specific rather than uniformly induced under chronic inflammation [14, 16].

ICOS-High Cells Exhibit Regulatory Bias Rather Than Pure Terminal Exhaustion

Functional module scoring highlighted significantly elevated regulatory T-cell–associated transcriptional programs in $ICOS^{high}$ cells. These include FOXP3, IL2RA, CTLA4, and TNFRSF18. While checkpoint-associated exhaustion markers (PDCD1, TIGIT, HAVCR2) were likewise enriched, the magnitude of regulatory coupling exceeded exhaustion coupling. CD8-restricted analysis did not reveal significant differences in exhaustion within $CD8^+$ subsets alone. These outcomes indicate that ICOS-associated transcriptional reprogramming is not simply a sign of classical CD8 terminal exhaustion. Rather, it shows a more extensive immunoregulatory architecture [9, 10].

These observations are consistent with earlier evidence that shows that ICOS signaling supports Treg stability and suppressive capacity via PI3K–AKT pathways [14, 15, 17]. In tumor settings, ICOS expression plays a double-faced role, either as a marker of effector activation or as a marker of regulatory T cells expansion [16, 17]. Our data suggest that in HBV-associated HCC, $ICOS^{high}$ cells occupy a hybrid regulatory–activation state instead of representing purely terminally exhausted cytotoxic T cells.

It is notable that differential expression analysis identified 4,324 of 4,395 tested genes as statistically significant. Given the large cell numbers and power inherent to single-cell datasets, widespread statistical significance is not unexpected [19]. However, effect sizes varied substantially, and the strongest enrichment signals were concentrated within regulatory and checkpoint-associated programs rather than indiscriminately across all gene classes. Thus, the biological interpretation rests not on the breadth of statistical signals but on coherent, functionally meaningful enrichment patterns.

Hallmark Pathway Analysis Shows Relative Reduction of Tumor-Conditioned Programs in ICOS-High Cells

This study's analysis identified 30 pathways that were significantly enriched and more active in ICOS-low. This shows a greater engagement of these pathways in $ICOS^{low}$ cells than in $ICOS^{high}$cells. Instead of assuming complete suppression, these results suggest that $ICOS^{low}$ cells have greater engagement with tumor-associated stress, metabolic remodeling, and stromal interaction programs [6, 12].

Pathways such as Hallmark TGFB Signalling and Hallmark Epithelial to Mesenchymal transition were much more active in $ICOS^{low}$ cells. Given the central role of TGF-β in tumor-induced immunosuppression and immune exclusion, these observations concur with well-known mechanisms of tumor immune remodeling [4, 6]. The reduction of these signatures in $ICOS^{high}$ cells may show a partial shield from tumor-driven transcriptional conditioning. Enrichment of xenobiotic metabolism and coagulation

pathways in ICOS^{low} cells suggests greater integration of microenvironmental stress responses in that compartment.

The absence of positively enriched Hallmark programs in ICOS^{high} cells indicates that ICOS identity is not defined by dominant activation of an inflammatory or proliferative pathway. Rather, ICOS^{high} cells appear transcriptionally specialized rather than globally hyperactivated, consistent with a regulatory-biased immune state [14, 16].

Trajectory Modeling Repositions ICOS as an Early Differentiation Marker

Diffusion pseudotime analysis gives vital mechanistic context. ICOS^{high} cells were significantly enriched at early stages of differentiation. We observed a 61.5% lower median pseudotime compared to ICOS^{low} cells. Moreover, exhaustion score declined progressively along pseudotime ($\rho = -0.3601$), indicating that checkpoint-associated transcriptional intensity was highest in early differentiation states and decreased toward later positions.

This data contradict linear models in which exhaustion monotonically accumulates during differentiation [9, 10]. Instead, our data support a model in which early antigen engagement induces checkpoint expression as part of physiological activation feedback, followed by transcriptional remodeling along the differentiation continuum. ICOS expression, therefore, appears linked to early activation dynamics rather than terminal dysfunction [14].

The dynamic distribution of ICOS^{high} proportions across pseudotime bins confirms our suspicion. Early enrichment (14–15%) declines in later states (~ 3%), consistent with transient induction rather than stable lineage commitment. ICOS may therefore function as a temporal activation marker within chronic viral tumour microenvironments [6].

Bulk Validation Establishes a Tumor-Level ICOS–Exhaustion Axis

At the tumor scale, TCGA-LIHC bulk RNA-seq analysis exhibited strong monotonic coupling between ICOS and the exhaustion score, as well as between CTLA4 and PDCD1. Multivariable regression confirmed ICOS as the strongest independent predictor of exhaustion intensity, even after adjustment for cytotoxic infiltration. The results of the mediation analysis showed that about one-third of the ICOS effect relied on CD8 abundance, while a substantial direct effect persisted.

These results suggest that ICOS functions within a coordinated immune regulatory network rather than acting solely as a surrogate for cytotoxic infiltration [6, 16].

The relationship between ICOS and CD8 further supports cooperative amplification within the tumor microenvironment.

ICOS did not predict overall survival independently. This was despite the observed strong immune coupling. This substantiates the notion that ICOS reflects immune-state architecture rather than functioning as a standalone prognostic biomarker in untreated cohorts [8, 18].

Limitations

Trajectory inference relies on computational prediction and assumes continuity throughout differentiation space [20]. Although diffusion mapping is widely used, pseudotime does not infer causality. Second, transcriptional signatures were not validated at the protein level. Third, functional assays assessing suppressive capacity, cytokine production, or proliferative potential were not available.

Conclusion

This integrative analysis redefines ICOS expression within HBV-associated HCC as a marker of early-positioned, regulatory-biased immune activation rather than terminal exhaustion. $ICOS^{high}$ cells occupy early differentiation states, demonstrating coherent regulatory transcriptional programming. They also show relative depletion of tumor-conditioning pathways and show strong tumor-level coupling to exhaustion intensity. However, ICOS does not independently dictate survival in untreated tumors. These results place ICOS as an immune regulatory hub coordinating activation, regulation, and exhaustion dynamics in the tumor microenvironment. By combining single-cell resolution with bulk validation, this study provides mechanistic clarity on ICOS biology in chronic viral malignancy and informs rational therapeutic methods targeting the ICOS axis.

Declarations

Ethical Approval:

Ethical approval was not required for this study, as all analyses were performed using publicly available, de-identified datasets obtained from the Gene Expression Omnibus (GEO) database.

Informed Consent:

Informed consent was not required for this study, as it involved secondary analysis of publicly available, anonymised data.

Funding:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions:

All authors contributed equally to the conception, design, and execution of the study. The overall manuscript was drafted and critically revised by all authors for important intellectual content. All authors read and approved the final manuscript.

Data Availability:

The datasets analysed during the current study are publicly available in the Gene Expression Omnibus (GEO) repository under accession numbers GSE89632 and GSE164760. All data used in this study can be accessed without restriction.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin 71(3):209–249. https://doi.org/10.3322/caac.21660
2. McGlynn KA, Petrick JL, El-Serag HB (2021) Epidemiology of hepatocellular carcinoma. Hepatology 73(1):4–13. https://doi.org/10.1002/hep.31288
3. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR (2019) A global view of hepatocellular carcinoma: trends, risk, prevention and management. Nat reviews Gastroenterol Hepatol 16(10):589–604. https://doi.org/10.1038/s41575-019-0186-y
4. Ringelhan M, Pfister D, O'Connor T et al (2018) The immunology of hepatocellular carcinoma. Nat Immunol 19:222–232. https://doi.org/10.1038/s41590-018-0044-z
5. Zhang, Q., He, Y., Luo, N., Patel, S. J., Han, Y., Gao, R., … Zhang, Z. (2019). Landscape and dynamics of single immune cells in hepatocellular carcinoma. Cell, 179(4), 829-845.10.1016/j.cell.2019.10.003
6. Shen KY, Zhu Y, Xie SZ et al (2024) Immunosuppressive tumor microenvironment and immunotherapy of hepatocellular carcinoma: current status and prospectives. J Hematol Oncol 17:25. https://doi.org/10.1186/s13045-024-01549-2
7. Finn, R. S., Qin, S., Ikeda, M., Galle, P. R., Ducreux, M., Kim, T. Y., … Cheng, A.L. (2020). Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma.New England Journal of Medicine, 382(20), 1894–1905.DOI: 10.1056/NEJMoa1915745
8. Llovet J, Zucman-Rossi J, Pikarsky E et al (2016) Hepatocellular carcinoma. Nat Rev Dis Primers 2:16018. https://doi.org/10.1038/nrdp.2016.18
9. Blank CU, Haining WN, Held W et al (2019) Defining 'T cell exhaustion'. Nat Rev Immunol 19:665–674. https://doi.org/10.1038/s41577-019-0221-9
10. Beltra, J. C., Manne, S., Hakeem, M. S. A., Kurachi, M., Giles, J. R., Chen, Z., …Wherry, E. J. (2020). Developmental relationships of four exhausted CD8 T cell subsets reveals underlying transcriptional

and epigenetic control mechanisms. The Journal of Immunology, 204(1_Supplement), 77 – 16.https://doi.org/10.4049/jimmunol.204.Supp.77.16

11. Ma, L., Hernandez, M. O., Zhao, Y., Mehta, M., Tran, B., Kelly, M., … Wang, X. W.(2019). Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer.Cancer cell, 36(4), 418-430.10.1016/j.ccell.2019.08.007
12. Sun, Y., Wu, L., Zhong, Y. U., Zhou, K., Hou, Y., Wang, Z., … Fan, J. (2021). Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. Cell, 184(2), 404–421.:10.1016/j.cell.2020.11.041
13. Wikenheiser DJ, Stumhofer JS (2016) ICOS Co-Stimulation: Friend or Foe? Front Immunol 7:304. 10.3389/fimmu.2016.00304
14. Dong C, Juedes A, Temann UA et al (2001) ICOS co-stimulatory receptor is essential for T-cell activation and function. Nature 409:97–101. https://doi.org/10.1038/35051100
15. Amatore F, Gorvel L, Olive D (2020) Role of Inducible Co-Stimulator (ICOS) in cancer immunotherapy. Expert Opin Biol Ther 20(2):141–150. https://doi.org/10.1080/14712598.2020.1693540
16. Nikanjam, M., Kato, S., Nishizaki, D., Miyashita, H., Pabla, S., Nesline, M. K., …Kurzrock, R. (2026). Inducible T-Cell Co-Stimulator (ICOS) and ICOS Ligand: Dealing With a Two-Faced Cancer Immunoregulatory System. Cancer medicine, 15(1), e71467. https://doi.org/10.1002/cam4.71467
17. Pratama A, Schnell A, Mathis D, Benoist C (2019) Developmental and cellular age direct conversion of CD4 + T cells into RORγ + or Helios+ colon Treg cells. J Exp Med 217(1):e20190428. https://doi.org/10.1084/jem.20190428
18. Sade-Feldman, M., Yizhak, K., Bjorgaard, S. L., Ray, J. P., de Boer, C. G., Jenkins,R. W., … Hacohen, N. (2019). Defining T cell states associated with response to checkpoint immunotherapy in melanoma. Cell, 176(1), 404. 10.1016/j.cell.2018.12.034
19. Zhang, Q., He, Y., Luo, N., Patel, S. J., Han, Y., Gao, R., … Zhang, Z. (2019). Landscape and dynamics of single immune cells in hepatocellular carcinoma. Cell, 179(4), 829–845.10.1016/j.cell.2019.10.003
20. Saelens W, Cannoodt R, Todorov H et al (2019) A comparison of single-cell trajectory inference methods. Nat Biotechnol 37:547–554. https://doi.org/10.1038/s41587-019-0071-9

Figures

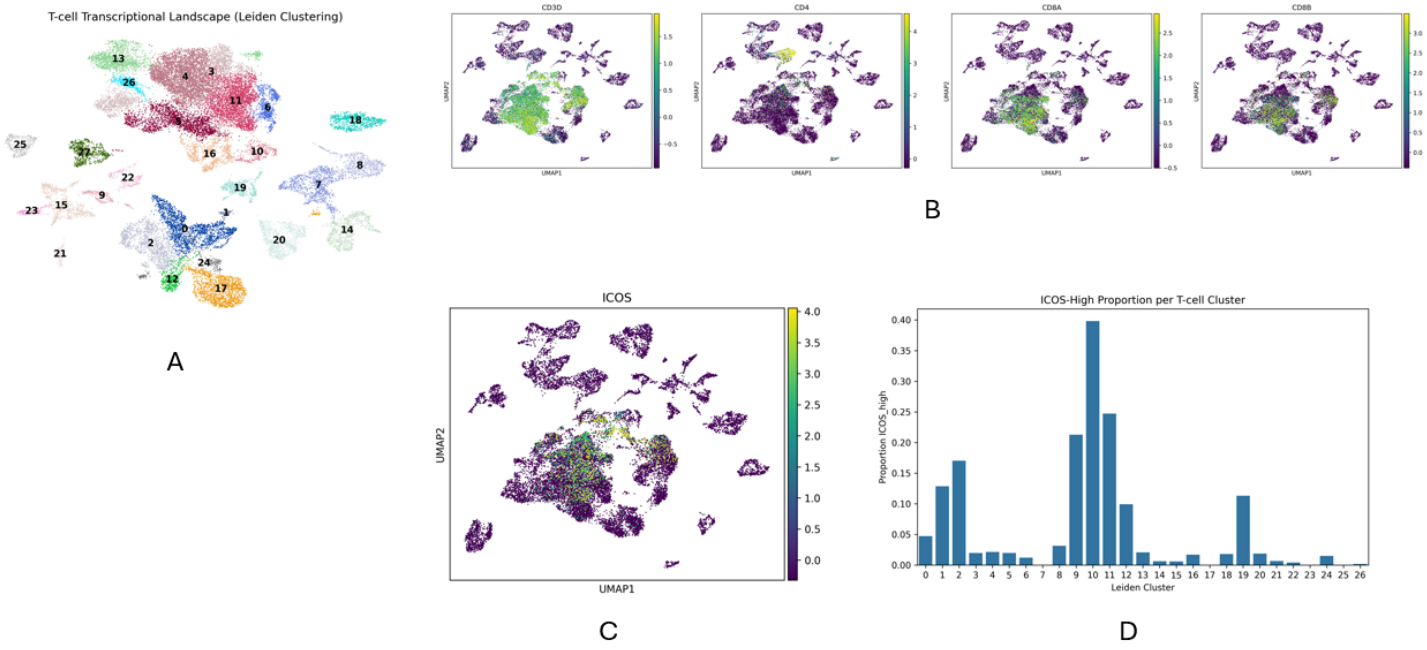


Figure 1

Transcriptional architecture of tumor-infiltrating T cells and spatial distribution of ICOS expression.

(A) Uniform Manifold Approximation and Projection (UMAP) of 25,590 quality-controlled tumor-infiltrating T cells following removal of contaminant clusters. Cells are colored by Leiden cluster identity (28 clusters), defining the transcriptional landscape used for all downstream analyses.

(B) UMAP feature plots showing canonical lineage markers (CD3D, CD4, CD8A, CD8B) confirming T-cell identity and delineation of $CD4^{+}$ and $CD8^{+}$ compartments. Spatial segregation of lineage-defining transcripts validates biological coherence of clustering.

(C) UMAP visualization of ICOS transcript expression across the T-cell landscape. ICOS expression is heterogeneously distributed and enriched within discrete transcriptional compartments.

(D) Quantification of $ICOS^{high}$ cell proportion across Leiden clusters. Bars represent the fraction of $ICOS^{high}$ cells per cluster, identifying specific T-cell subsets with preferential ICOS enrichment.

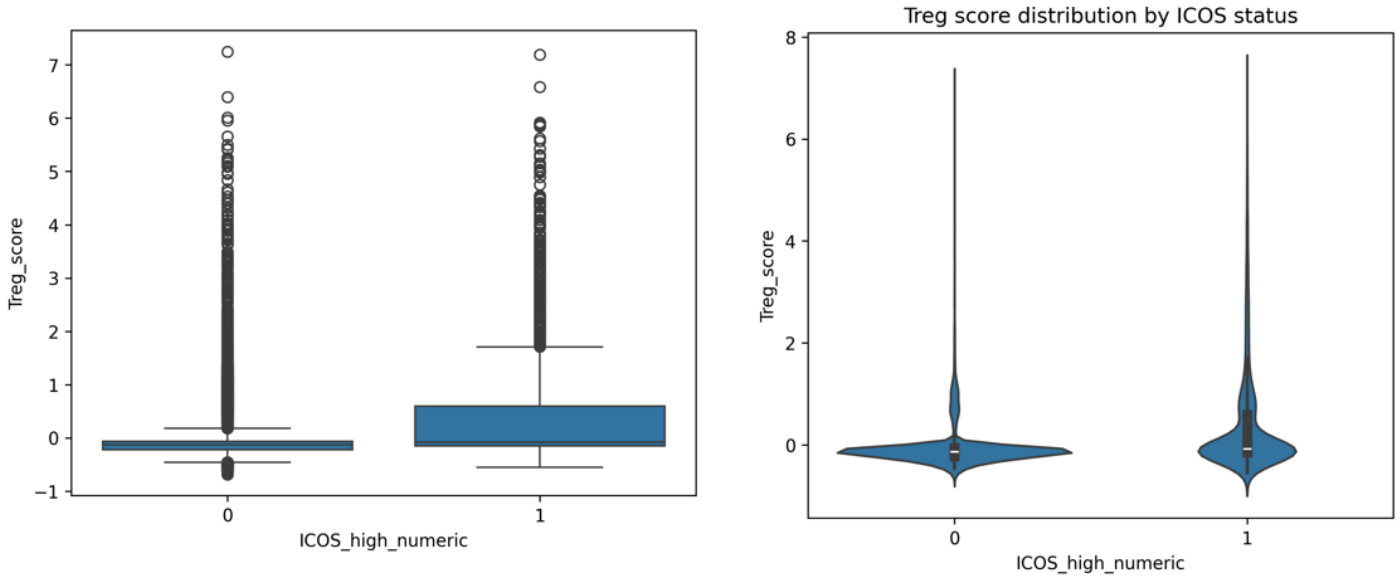


Figure 2

ICOS expression is associated with enrichment of regulatory T-cell transcriptional programs.

(A) Boxplot comparing Treg module scores between $ICOS^{high}$ and $ICOS^{low}$ T cells. $ICOS^{high}$ cells exhibit significantly elevated regulatory program activity, indicating transcriptional convergence with Treg-associated signatures.

(B) Violin plot illustrating the distributional shift of Treg scores by ICOS status. $ICOS^{high}$ cells display broader and upward-shifted score distributions, confirming enrichment beyond median effects and supporting functional polarization toward a regulatory phenotype.

Statistical comparisons were performed using non-parametric.

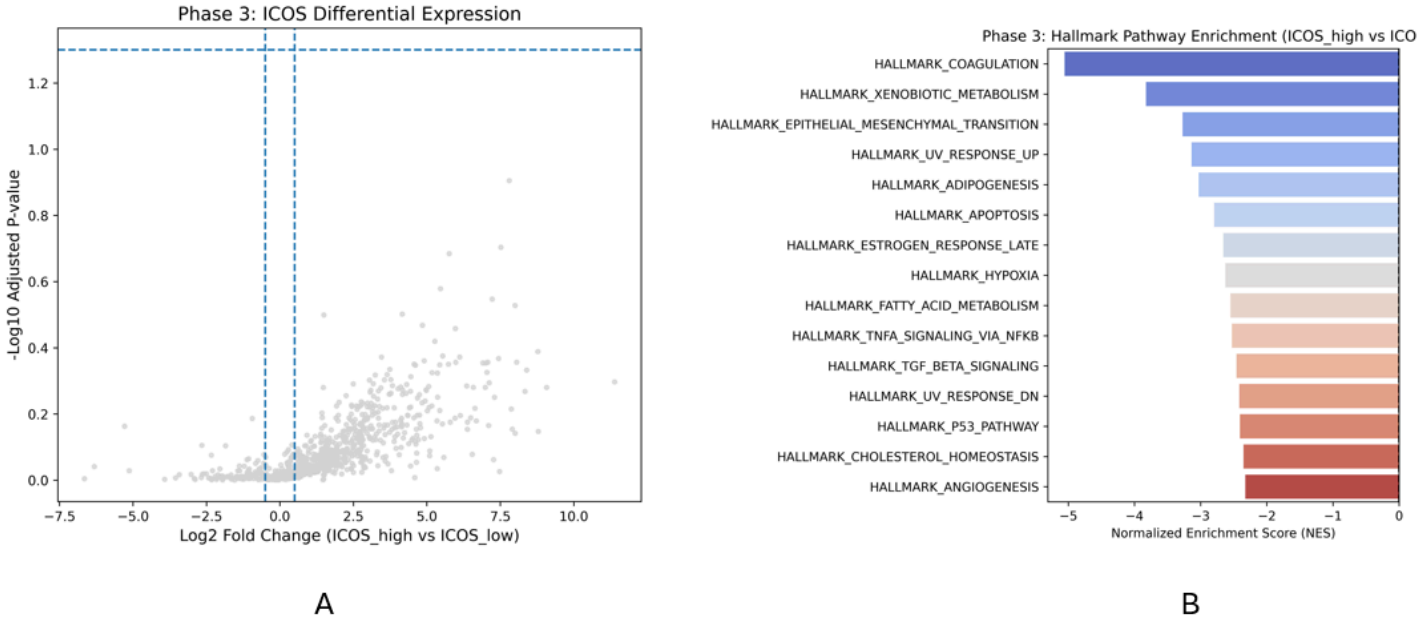


Figure 3

Differential transcriptional and pathway programs associated with ICOS-high T cells.

(A) Volcano plot of differential gene expression comparing$ICOS^{high}$ versus $ICOS^{low}$ T cells. Each point represents a gene plotted by log2 fold change and adjusted p-value. $ICOS^{high}$ cells demonstrate broad transcriptional reprogramming characterized by coordinated upregulation and downregulation of immune and metabolic genes.

(B) Hallmark Gene Set Enrichment Analysis (GSEA) of pre-ranked differential expression results. Bars represent normalized enrichment scores (NES). $ICOS^{high}$ cells exhibit significant enrichment of pathways associated with immune signaling, metabolic reprogramming, and stress-response programs, indicating functional remodeling linked to ICOS expression.

All pathway statistics and ranked gene lists are provided in Supplementary Tables.

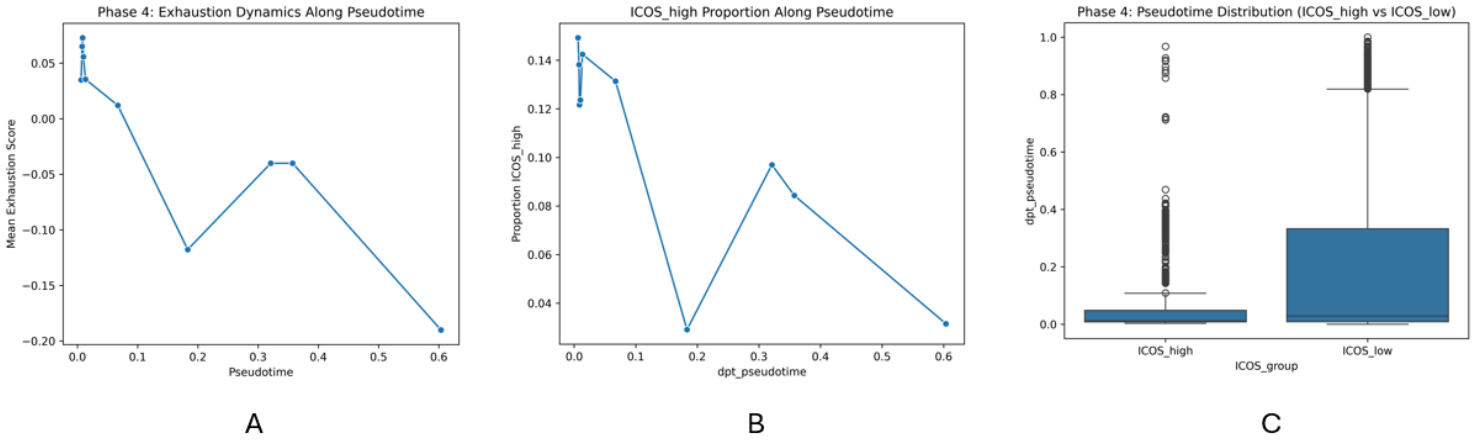


Figure 4

ICOS expression and exhaustion dynamics along diffusion pseudotime.

(A) Mean exhaustion score plotted across diffusion pseudotime bins, demonstrating dynamic regulation of exhaustion programs during T-cell state transitions.

(B) Proportion of $ICOS^{high}$ cells across pseudotime trajectory. ICOS enrichment is observed preferentially in early-to-intermediate pseudotime states, suggesting stage-specific regulatory involvement.

(C) Boxplot comparing pseudotime distributions between and $ICOS^{low}$ cells. $ICOS^{high}$ cells are significantly biased toward distinct trajectory regions, supporting a temporal relationship between ICOS expression and T-cell differentiation states.

These data position ICOS within the dynamic continuum of T-cell functional states.

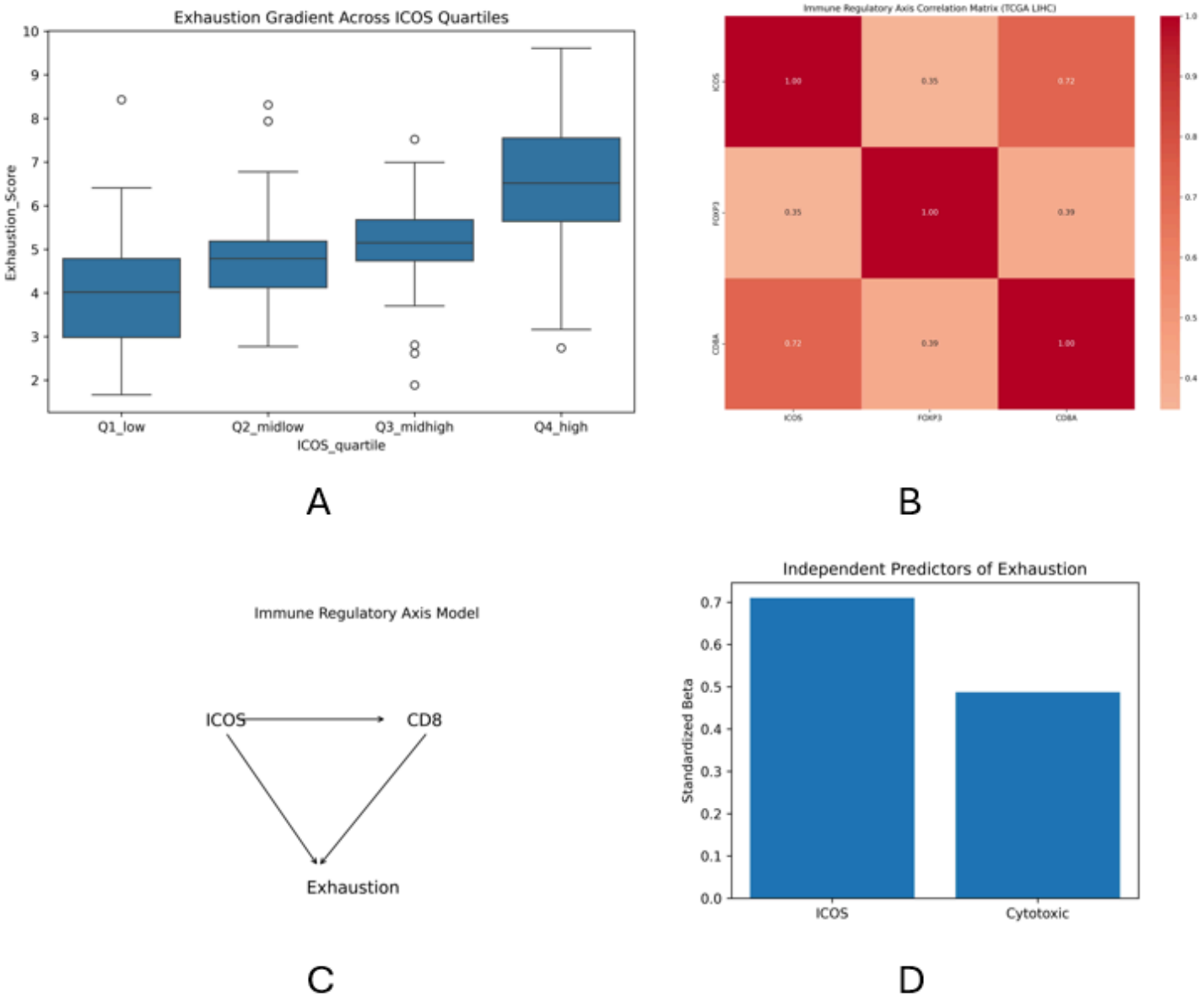


Figure 5

ICOS as a Component of an Immune Regulatory Axis (TCGA LIHC)

Figure 5. Systems-level modeling of the ICOS-centered immune regulatory axis in bulk tumor transcriptomes.

(A) Exhaustion score across quartiles of ICOS expression in TCGA LIHC. A stepwise increase in exhaustion score is observed across ICOS expression quartiles, demonstrating a dose-dependent relationship.

(B) Spearman correlation matrix among ICOS, FOXP3, and CD8A expressions. ICOS positively correlates with both regulatory (FOXP3) and cytotoxic (CD8A) markers, defining an integrated immune regulatory axis.

(C) Conceptual immune regulatory path model illustrating relationships among ICOS, CD8, and exhaustion programs.

(D) Multivariable regression model showing standardized beta coefficients for predictors of exhaustion. ICOS remains an independent predictor of exhaustion after adjustment for cytotoxic activity, supporting its regulatory contribution beyond simple T-cell abundance.

Full regression outputs and mediation analyses are provided in Supplementary Tables.

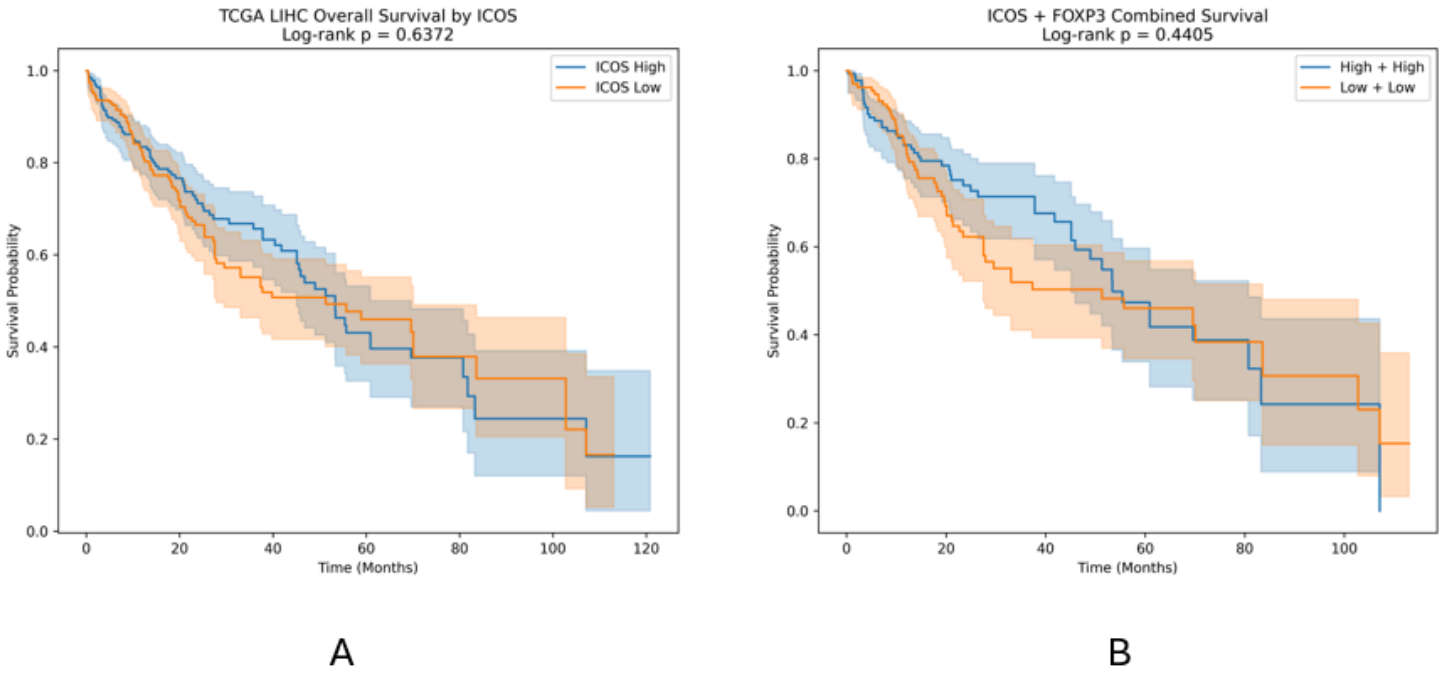


Figure 6

Clinical Relevance of ICOS in TCGA LIHC

Figure 6. Survival impact of ICOS expression in hepatocellular carcinoma (TCGA LIHC).

(A) Kaplan–Meier overall survival analysis stratified by ICOS expression. No statistically significant survival difference is observed between $ICOS^{high}$ and $ICOS^{low}$ groups (log-rank test).

(B) Combined ICOS and FOXP3 stratified survival analysis comparing high/high versus low/low groups. Despite transcriptional regulatory enrichment, combined immune regulatory signatures do not independently predict overall survival.

These results suggest that while ICOS shapes immune transcriptional programs, its prognostic impact is context-dependent.

Supplementary Files

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

- S0MasterIndex.xlsx
- S1ClusterLevelICOShighProportions.csv
- S2Phase2ICOSproportionsbylineage.csv
- S3FullDifferentialExpressionTable.csv
- S4RankedGeneListUsedforGSEA.csv
- S5SignificantHallmarkPathwaysFDR0.05.csv
- S6Top50UpDownGenes.csv
- S7DiffusionPseudotimeDistributionStatistics.csv
- S8FullHallmarkGSEAResults.csv
- S9BulkImmuneCorrelationMatrix.csv
- S10TCGAICOSFOXP3coxregression.csv
- S11InteractionModelOutputICOSCD8.txt
- S12MediationAnalysisICOSCD8Exhaustion.csv
- S13ICOSThresholdSensitivityAnalysis.csv
- S14ICOSThresholdEffectSizes.csv
- S15CoxRegressionResultsICOSFOXP3.csv