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.

Single-Cell Transcriptomic Profiling Reveals T-Cell Functional Shifts in Primary Versus Metastatic Colorectal Cancer

Muhammad Usman Saleem
International Islamic University, Islamabad https://orcid.org/0009-0007-9726-4414
Muhammad Imran Shabbir
imran.shabbir@iiu.edu.pk

International Islamic University, Islamabad
Ashar Ahmed
International Islamic University, Islamabad
Fahim Karamat
International Islamic University, Islamabad
Muhammad Faisal
International Islamic University, Islamabad
Abdul Muiz Ali
International Islamic University, Islamabad

Research Article

Keywords: colorectal cancer, liver metastasis, tumor immune microenvironment, T-cells, immune landscape, T-cell differentiation, single-cell transcriptomics

Posted Date: April 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9411430/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

Colorectal cancer (CRC) develops in the colon or rectum and is the second leading cause of cancer-related mortalities, with high rates of liver metastasis (LM). The tumor immune microenvironment (TIME) of primary CRC and LM-CRC is immunosuppressive, leading to T-cell dysfunction. Hence, this study employed computational techniques to unravel the novel T-cell-specific biomarkers being dysregulated due to niche-specific cancer cell signaling. A single-cell RNA sequencing (scRNA-seq) analysis of primary and LM CRC was performed, followed by identifying differentially expressed genes (DEGs) of T-cells. Subsequently, excluding common genes among both niches (primary CRC and LM-CRC), unique genes of T-cells were selected, and enriched Reactome pathways were identified. Moreover, the top-5 hub-genes of T-cells were identified by mapping the interactions of pathway genes in both niches. Lastly, novel hub-genes were selected after their expression validation in independent scRNA-seq datasets. The novel hub-genes, being dysregulated due to distinct cancer cell signals in respective niches, indicated to be implicated in the dysfunction of niche-specific T-cells, providing potential targets for immunotherapies. However, further larger cohorts are required, and clinical validation is essential to validate the therapeutic relevance of identified novel biomarkers.

1. Introduction

Colorectal cancer (CRC) is a heterogeneous and complex disease characterized by the accumulation of genetic and epigenetic alterations leading to uncontrolled growth of cells and metastasis (1). According to the American Cancer Society; CRC is a second leading cause of cancer related deaths; while an estimated 107,320 cases of colon cancer and 46,950 cases of rectal cancer will be diagnosed in the US, and 52,900 deaths due to colon and rectum cancers in 2025 (2).

Moreover, about 25% of CRC patients are initially diagnosed with distant metastasis and 50% will develop metastases at a later stage (3). However, liver metastasis (LM) of CRC is the most common route for distant spread in advanced stages and is the leading cause of death among CRC patients. Additionally, 20-25% of the CRC patients, at the time of diagnosis, exhibit LM; while 50% of the patients may develop metachronous LM following excision of the primary tumor (4).

Furthermore, LM poses significant challenges for treatment and remains a leading cause of mortality. Surgical resection and chemotherapy are the primary treatment options for CRC-LM; however, at the time of diagnosis only 10-20% of CRC-LM patients are deemed suitable for surgical intervention (5). Additionally, over the past two decades, chemotherapy, targeted therapies, and immunotherapies have improved the treatment of metastatic CRC; however, despite the advancements in treatment options, it remains a major cause of cancer-related mortalities worldwide (6).

CRC-LM exhibits a decreased anti-tumor immunity and efficacy of immunotherapy (7). This is due to dynamic interactions of immune and tumor cells within the tumor immune microenvironment (TIME), reshaping immune responses and influencing therapeutic outcomes (8). In CRC, immune-related pathways serve as critical regulators of both tumor biology and TIME; while dysregulation of such

pathways through aberrant activation or suppression disrupts the immune cell populations and their functionality, eventually leading to immune evasion mechanisms supporting tumor growth and metastasis (9).

Recruitment of immune cells such as regulatory T-cells (Tregs) and myeloid-derived suppressor cells (MDSCs) is observed in advanced CRC, leading to the suppression of anti-tumor immune responses and promoting tumor progression. Moreover, immunosuppressive cytokines such as TGF-β and interleukin-10 (IL-10) are produced by Tregs to suppress cytotoxic T-cell response and maintain immune tolerance, reflecting a critical role of immune cells in TIME of primary and metastatic CRC (10).

Moreover, $CD8^+$ T-cells, $CD4^+$ T-cells, and their subsets play a major role in the anti-tumor activity within TIME; however, in CRC, effective T-cell-mediated immunity is opposed at several stages such as exclusion of cytotoxic $CD8^+$, γδ and Th1 cells from the lesion. The T-cells which do infiltrate become anergic or drift into terminal exhaustion marked by multiple inhibitory checkpoints and loss of cytotoxicity (11).

Several single-cell RNA sequencing (scRNA-seq) studies investigated the role of immune cells, specifically T-cells within the TIME. A scRNA-seq study explored the $CD4^+$ T-cells within the CRC TIME showing the expression of three genes, including *LAYN, MAGEH1* and *CCR8* of Tregs being influenced by TIME. Another study explored the cell heterogeneity and dynamic relationship between $CD4^+$ and $CD8^+$ T-cell subpopulations through integrated transcriptomic analysis (12).

Conclusively, the immune landscape within the TIME of CRC and LM-CRC plays a critical role in reshaping therapeutic landscape and high throughput technologies allow to explore the heterogeneity of immune cells at single cell resolution level. Therefore, this study aimed to unravel the complex interplay between niche-specific tumor signaling and T-cell dysfunction, highlighting how primary and metastatic tumors differentially modulate T-cell responses within their respective microenvironments.

This was achieved by using computational techniques such as scRNA-seq analysis, followed by gene expression profiling, identification of enriched pathways, and mapping gene interactions. These analyses led to the identification of novel biomarkers of T-cells, which were further validated using independent scRNA-seq datasets. Hence, by broadening the understanding of the immune landscape, this study provides potential novel targets for improved immunotherapies. However, further clinical investigations are required to validate the findings of this study and assess their therapeutic relevance.

2. Methodology

2.1. Retrieval of scRNA-seq data

The scRNA-seq preprocessed (mapped) dataset of CRC with an accession ID of GSE231559 was retrieved using the Gene Expression Omnibus (GEO) datasets (https://www.ncbi.nlm.nih.gov/gds/) of National Center for Biotechnology Information (NCBI), an open-source database providing functional

genomics studies and their data (13). The dataset consisted of four conditions, including CRC-Tumor, CRC-Normal, LM-Tumor, and LM-Normal. Among these conditions, 6 samples were of CRC-Tumor; 3 of CRC-Normal; 9 of LM-Tumor; and 8 of LM-Normal, comprising a total of 26 samples taken from CRC patients, which were sequenced using Illumina NextSeq 500 platform.

Additionally, the “Seurat” package of R was utilized to combine the aforementioned retrieved samples of the CRC dataset using a “merge” function that resulted in a combined matrix. Moreover, the combined matrix was converted into a “Seurat Object” using the “CreateSeuratObject” function, for further preprocessing of the dataset.

2.2. Preprocessing and cell clustering

Initially, the dataset exhibited 38224 genes and 133966 cells prior to data filtration. Subsequently, the combined “Seurat Object” was subjected to preprocessing steps to filter out low-quality and dead cells. The cells exhibiting genes between 200 and 6000, unique molecular identifiers (UMIs)/RNA counts greater than 1000, and mitochondrial reads of less than 10% were retained for further analysis.

Furthermore, normalization of the filtered dataset was performed by utilizing the “LogNormalize” method with a scale factor of “10000”, followed by the identification of the 2000 highly variable features (HVFs) through the “vst” method. Moreover, data scaling followed by the principal component analysis (PCA) for the dimensionality reduction of the dataset were performed. Lastly, functions such as “FindNeighbors”, “FindClusters”, along with 20 principal components (PCs) were used for data clustering at a resolution of 0.5 by utilizing the Leiden algorithm.

2.3. Identification of cell types

Cell type annotation of the cell clusters was performed using the ScType database (https://sctype.app/), an open-source automated computational platform providing clinically validated cell-type identification of the scRNA/snRNA-seq datasets (14). Subsequently, immune system specific positive and negative marker gene sets were prepared by utilizing the “gene_set_prepare” function. The expression of prepared marker genes was used for annotation by computing the cell-type-specific scores through the “sctype_score” function. Moreover, aggregation of the cell-type-specific scores at cluster level was performed to identify the most likely cell type associated with each cluster.

2.4. Gene expression profiling of T-cell subsets in primary CRC and LM-CRC

The differential gene expression (DGE) analysis of identified T-cells, including $CD8^+$ NKT-like cells, memory $CD4^+$ T-cells and naive $CD4^+$ T-cells was performed between diseased and normal conditions

(CRC-Tumor vs CRC-Normal and LM-Tumor vs LM-Normal) by using “FindMarkers” function with a threshold of ± 1 and “wilcox” method. This resulted in the dysregulated genes of T-cells within the TIME in both groups and the resultant p-values were subjected to false discovery rate (FDR) corrections through the Benjamini-Hochberg (BH) method by utilizing the “p.adjust” function.

Subsequently, the upregulated genes of T-cells were further selected and compared with respective T-cell from both groups, such that the differentially expressed genes (DEGs) of $CD8^+$ NKT-like cells from “CRC-Tumor vs CRC-Normal” were compared with the DEGs of $CD8^+$ NKT-like cells from “LM-Tumor vs LM-Normal” group; similarly other T-cells (memory $CD4^+$ T-cells and naive $CD4^+$ T-cells) were compared in both groups with the same approach. This was achieved by using the Interactivenn (https://www.interactivenn.net/), a web-based tool that allows users to analyze sets of data using Venn diagrams (15).

Moreover, the unique genes of $CD8^+$ NKT-like cells in “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups were selected; similarly the unique genes of memory $CD4^+$ T-cells and naive $CD4^+$ T-cells in both groups were selected for further analysis. This was performed to gain insights into the diverse roles of primary and LM cancer cells of CRC in differentially modulating T-cells within TIME for their survival, growth, and proliferation.

2.5. Enriched T-cells pathways and gene interactions mapping

The selected unique genes of T-cells among both groups were subjected to pathways enrichment analysis using the Python package of Gene Set Enrichment Analysis (GSEA/GSEApy) (https://github.com/zqfang/GSEApy), an open-source computational method to analyze gene expression data and identify significant pathways or biological processes enriched in biological states (16). The “Reactome_Pathways_2024” gene-set library of Enrichr (identify enriched biological processes, pathways, and other relevant categories by analyzing lists of genes, https://maayanlab.cloud/Enrichr/) interface from GSEApy was utilized to analyze the enriched pathways of identified DEGs of T-cells.

Subsequently, the resultant significant enriched pathways were selected and their genes were used to perform protein-protein interaction (PPI) analysis using Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (https://string-db.org/), which systematically collects and integrates PPIs, both physical interactions as well as functional associations (17).

Lastly, the top-5 hub-genes of each T-cells were identified by utilizing the protein networks resulting from STRING. This was achieved by using the Cytoscape, an open-source software integrating the biomolecular interaction networks with high-throughput interaction data (18).

2.6. Expression validation of candidate T-cells hub-genes

The identified top-5 hub-genes of $CD8^+$ NKT-like cells, memory $CD4^+$ T-cells and naive $CD4^+$ T-cells were subjected to the expression validation using the Tumor Immune Single Cell Hub 2 (TISCH2) (http://tisch.comp-genomics.org/), an open-source scRNA-seq data resource comprising human and mouse tumor datasets, providing extensive characterization of gene expressions in tumor microenvironment (19).

Subsequently, two datasets of CRC from TISCH2 were used, where one dataset with an accession ID of GSE166555 consisted of primary and normal CRC states with 66,050 cells; while other dataset with an accession ID of GSE136394 comprised metastatic state with 67,171 cells. Notably, the dataset (GSE136394) comprising metastatic state consisted of treated (immunotherapy) and untreated samples.

Moreover, the gene expression matrix from both datasets were retrieved along with their cell types metadata and heatmaps were generated for both datasets. The hub-genes of T-cells from primary state cancer were validated in the GSE166555 dataset, while the hub-genes from LM state were validated in the GSE136394 dataset. This was performed to validate the expression levels of top-5 hub-genes in T-cells of other independent scRNA-seq datasets.

3. Results

3.1. Data preprocessing and annotated cell clusters

The quality control measures retained 38224 genes and 59229 cells in the dataset; while normalization and scaling were performed and HVFs were identified. Among the 2000, the top-10 HVFs included *CCL4L2, S100A8, C1QB, CXCL8, IL1B, G0S2, C1QA, IFI27, GZMB,* and *C1QC*. Subsequently, the linear dimensionality reduction and cell clustering were performed based on the gene expression profiles, exhibiting a total 22 cell clusters. The data filtration violin plots, HVFs plot, PCA plot, elbow plot, and cell clusters are shown in **Supplementary Figure S1 (A-E)**.

Moreover, the cell type annotation resulted in 15 cell types within the TIME of primary and LM CRC, including memory $CD4^+$ T cells, endothelial, pre-B cells, cancer cells, myeloid dendritic cells, plasma B cells, macrophages, naive $CD4^+$ T cells, plasmacytoid dendritic cells, classical monocytes, $CD8^+$ NKT-like cells, platelets, effector $CD8^+$ T cells, non-classical monocytes, and natural killer cells. The annotated cell clusters are shown in **Figure 1 (A)**.

Furthermore, it was observed that the data filtration process filtered out the majority of the cell types from CRC-Normal condition, while it also exhibited a small number of cancer cells indicating a metastasis of cancer cells. However, it was observed that naive $CD4^+$ T-cells and memory $CD4^+$ T-cells were found in majority as compared to $CD8^+$ NKT-like cells and effector $CD8^+$ T-cells in CRC-Tumor

condition, indicating a non-cytotoxic immune landscape, likely reflecting immunosuppressive signaling and immune evasion.

Additionally, a slight increase in $CD8^+$ NKT-like cells were observed in LM-Tumor compared to LM-Normal condition; while major increase in naive $CD4^+$ T-cells, memory $CD4^+$ T-cells, and effector $CD8^+$ T-cells in LM-Tumor compared to LM-Normal condition was observed, indicating broad lymphocyte recruitment; however, the large number of naive $CD4^+$ T-cells may suggest incomplete priming or an ongoing immune modulation by metastatic cancer cells.

Notably, there were a total of four effector $CD8^+$ T-cells in CRC-Normal condition; therefore these T-cells were excluded from all four conditions for further analysis as they would not yield significant DGE results and might lead to non-significant conclusions. The annotated cell types in CRC-Tumor, CRC-Normal, LM-Tumor, and LM-Normal conditions are shown in **Supplementary Figure S1 (F)**.

3.2. Gene expression profiling of T-cell subsets in primary CRC and LM-CRC

The DGE analysis of $CD8^+$ NKT-like cells, memory $CD4^+$ T-cells and naive $CD4^+$ T-cells of the "CRC-Tumor vs CRC-Normal" group exhibited 447, 698, and 126 upregulated genes, respectively. While, the top-10 upregulated genes of $CD8^+$ NKT-like cells included *XIST, LSR, SLC30A1, GRAMD1B, SLC4A10, SNX9, PMEPA1, GFPT2, CXCL8* and *CTLA4* with log2FC values of 7.72, 5.42, 4.53, 4.38, 4.16, 4.04, 3.84, 3.59, 3.56 and 3.53; and FDR adjusted p-values of 1.93E-12, 1.88E-03, 3.18E-02, 2.56E-02, 4.37E-02, 2.15E-09, 2.24E-04, 1.77E-02, 1.81E-02 and 3.23E-04, respectively; whereas, memory $CD4^+$ T-cells exhibited *RTKN2, ANXA1, LAYN, TNFRSF9, GPR171, HPGD, IGHA1, CAPG, METRNL,* and *PDE4A* with log2FC values of 6.48, 6.44, 6.30, 6.17, 5.70, 5.64, 5.58, 5.13, 5.10, and 4.97; and FDR adjusted p-values of 7.83E-05, 1.50E-02, 1.80E-05, 1.52E-04, 1.93E-03, 6.51E-04, 4.32E-03, 2.40E-03, 3.64E-05, and 2.57E-03, respectively; lastly, naive $CD4^+$ T-cells exhibited *CEBPD, FAM129A, IGHA1, CCR6, BHLHE40, LMNA, ANKRD28, CTSH, FURIN,* and *LINC00513* with log2FC values of 4.91, 4.04, 3.34, 2.62, 2.49, 2.36, 2.34, 2.26, 2.23, and 2.22 and FDR adjusted p-values of 1.63E-07, 4.99E-14, 1.63E-07, 3.67E-09, 5.05E-06, 4.15E-08, 2.21E-08, 1.59E-04, 2.19E-05, and 1.11E-10, respectively. The volcano plots of each T-cell DEGs of "CRC-Tumor vs CRC-Normal" group are shown in **Figure 1 (B-D)**.

Moreover, the DGE of $CD8^+$ NKT-like cells, memory $CD4^+$ T-cells and naive $CD4^+$ T-cells of "LM-Tumor vs LM-Normal" group exhibited 70, 318, and 191 upregulated genes, respectively. Whereas, the top-10 upregulated genes of $CD8^+$ NKT-like cells included *NIBAN1, KRT86, LDLRAD4, PMEPA1, HIST1H3D, AL627171.2, SNHG32, FAM174C, SNHG1,* and *SNHG5* with log2FC values of 3.56, 3.18, 3.11, 2.64, 2.58, 2.46, 2.18, 2.18, 2.17, and 2.15 and FDR adjusted p-values of 5.72E-187, 9.33E-77, 5.19E-111, 7.06E-145, 2.57E-69, 4.21E-160, 4.52E-58, 6.37E-45, 5.98E-48, and 5.32E-141, respectively; while, memory $CD4^+$ T-cells exhibited *CXCL13, TSPAN13, LINC-PINT, GNG10, C17orf67, WDR47, BOLA2, CIAO2A, NMB,* and

GABPB1-IT1 with log2FC values of 5.61, 5.52, 3.53, 3.49, 3.30, 3.29, 3.27, 3.17, 3.17, and 3.12; and FDR adjusted p-values of 9.81E-04, 2.58E-03, 1.24E-25, 2.58E-03, 5.91E-03, 7.81E-03, 3.85E-03, 3.24E-03, 5.32E-04, and 2.07E-04, respectively; lastly, the naive $CD4^+$ T-cells exhibited *CCL20, GNLY, BOLA2, ZFAND2A, PMEPA1, SNHG1, SNHG5, TNFSF13B, HSPH1,* and *FAM174C* with log2FC values of 2.76, 2.73, 2.62, 2.37, 2.27, 2.26, 2.23, 2.22, 2.21, and 2.20; and FDR adjusted p-values of 3.92E-47, 1.52E-18, 1.78E-22, 4.75E-39, 7.04E-120, 2.48E-43, 1.41E-217, 2.70E-20, 1.00E-262, and 2.81E-27, respectively. The volcano plots of each T-cell DEGs of "LM-Tumor vs LM-Normal" group are shown in **Figure 1 (E-G)**.

Furthermore, it was observed that in the "CRC-Tumor vs CRC-Normal" group, $CD8^+$ NKT-like cells exhibited a total 2149 upregulated genes which ranged between 0.1 and 7.72 log2FC values; while exhibited a total 1322 upregulated genes which ranged between 0.1 and 3.56 log2FC values in "LM-Tumor vs LM-Normal" group. Similarly, memory $CD4^+$ T-cells exhibited a total of 2629 (0.1-6.48 log2FC) and 2003 (0.1-5.61 log2FC) upregulated genes in "CRC-Tumor vs CRC-Normal" and "LM-Tumor vs LM-Normal" groups, respectively; lastly, naive $CD4^+$ T-cells exhibited a total 1928 (0.1-4.91 log2FC) and 1614 (0.1-2.76 log2FC) upregulated genes in "CRC-Tumor vs CRC-Normal" and "LM-Tumor vs LM-Normal" groups, respectively.

This indicated an intense transcriptional activation, suggesting a tumor-driven dysregulation of T-cells such as immune exhaustion, suppression, or reprogramming in primary CRC microenvironment. On the other hand, an immune-silent microenvironment was indicated in the LM condition where T-cells exhibited fewer upregulated genes compared to the primary CRC.

Additionally, common and unique DEGs of T-cell subsets were identified using Interactivenn (https://www.interactivenn.net/). This was performed to specifically identify the unique genes of T-cell subsets respective to their niche, dysregulated due to cancer-driven signals in TIME of both niches (primary and metastatic). The Venn diagrams showed 15 common genes of $CD8^+$ NKT-like cells in both groups (CRC-Tumor vs CRC-Normal and LM-Tumor vs LM-Normal); 432 unique to "CRC-Tumor vs CRC-Normal"; while 55 unique genes to "LM-Tumor vs LM-Normal" group. Moreover, 56 common genes of memory $CD4^+$ T-cells in both groups; 642 unique to "CRC-Tumor vs CRC-Normal"; and 262 unique genes to the "LM-Tumor vs LM-Normal" group. Lastly, only 12 common genes of naive $CD4^+$ T-cells in both groups; 114 unique to "CRC-Tumor vs CRC-Normal"; and 179 unique genes to "LM-Tumor vs LM-Normal" group.

Interestingly, it was observed that only a few genes were found common among both niches (CRC-Tumor vs CRC-Normal and LM-Tumor vs LM-Normal) of respective T-cells, indicating that the cancer cells within TIME of primary and metastatic states are actively involved in signalling to dysregulate the T-cells in distinct ways compared to each other, eventually leading to immune suppression and evasion.

Therefore, the unique genes of T-cells were selected, exhibiting a total 6 gene sets ($CD8^+$ NKT-like cells, memory $CD4^+$ T-cells and naive $CD4^+$ T-cells belonging to "CRC-Tumor vs CRC-Normal" and "LM-Tumor

vs LM-Normal" groups) for pathway enrichment analysis. The common and unique genes among both groups (CRC-Tumor vs CRC-Normal and LM-Tumor vs LM-Normal) are shown in **Figure 1 (H-J)**.

3.3. Enriched pathways leading to aberrant T-cells in primary CRC and LM-CRC

The "CRC-Tumor vs CRC-Normal" group showed enriched top-10 significant Reactome pathways of $CD8^+$ NKT-like cells to be "Cytokine Signaling in Immune System", "Immune System", "Gene Expression (Transcription)", "RNA Polymerase II Transcription", "Signaling by Interleukins", "Cellular Responses to Stimuli", "Generic Transcription Pathway", "FOXO-mediated Transcription", "DARPP-32 Events", and "Basigin Interactions"; while memory $CD4^+$ T-cells included "Cytokine Signaling in Immune System", "Immune System", "Signaling by Interleukins", "Costimulation by the CD28 Family", "Phosphorylation of CD3 and TCR Zeta Chains", "Signaling by FLT3 ITD and TKD Mutants", "Interferon Signaling", "MHC Class II Antigen Presentation", "TCR Signaling", and "Translocation of ZAP-70 to Immunological Synapse" top-10 pathways; lastly, naive $CD4^+$ T-cells exhibited only three significant Reactome pathways, including "Cytokine Signaling in Immune System", "Immune System", and "Signaling by Interleukins". The pathway, "Signaling by FLT3 ITD and TKD Mutants" (memory $CD4^+$ T-cells), was observed to be relevant to acute myeloid leukemia (AML) rather than CRC and was excluded from PPI analysis. The enriched Reactome pathways of "CRC-Tumor vs CRC-Normal" group T-cells are shown in **Supplementary Figure S2 (A-C)**.

Additionally, the aforementioned Reactome pathways of "CRC-Tumor vs CRC-Normal" group indicates transcriptional reprogramming, heightened immune signaling, and stress responses through tumor-derived signals. The chronic activation of such pathways reflects T-cell dysfunction, such as impaired cytotoxicity, exhaustion, and loss of proliferative capacity. Collectively, these aforementioned pathways reflect a maladaptive T-cell response due to cancer-associated signals, eventually leading to T-cell dysfunction and tumor immune escape.

Moreover, the "LM-Tumor vs LM-Normal" group exhibited top-10 Reactome pathways of memory $CD4^+$ T-cells to be "Attenuation Phase", "Cellular Responses to Stimuli", "HSF1 Activation", "Cellular Responses to Stress", "HSF1-dependent Transactivation", "Cellular Response to Heat Stress", "HSP90 Chaperone Cycle for Steroid Hormone Receptors (SHR) in the Presence of Ligand", "Regulation of HSF1-mediated Heat Shock Response", "Mitochondrial Unfolded Protein Response (UPRmt)", and "Transport of Mature mRNA Derived From an Intron-Containing Transcript"; while naive $CD4^+$ T-cells exhibited "Cellular Responses to Stimuli", "Cellular Responses to Stress", "HSP90 Chaperone Cycle for Steroid Hormone Receptors (SHR) in the Presence of Ligand", "Attenuation Phase", "HSF1 Activation", "HSF1-dependent Transactivation", "Cellular Response to Heat Stress", "Regulation of HSF1-mediated Heat Shock Response", "Selective Autophagy", and "Autophagy" top-10 Reactome pathways; lastly, it was observed that upregulated genes of $CD8^+$ NKT-like cells did not result in any significant enriched Reactome pathways, therefore these cells

were excluded from further analysis. The enriched Reactome pathways of “LM-Tumor vs LM-Normal” group T-cells are shown in **Supplementary Figure S2 (D-E)**.

Furthermore, the aforementioned Reactome pathways of “LM-Tumor vs LM-Normal” group suggested cellular stress responses, such as heat shock responses, mitochondrial unfolded protein response, and autophagy, which are hallmarks of cellular adaptations to proteotoxic and metabolic stress. This reflects the protective but non-functional state of T-cells in the liver niche due to the metastatic cancer-derived signals, imposing a stressful and suppressive microenvironment.

3.4. Candidate T-cells hub-genes implicated in primary CRC and LM-CRC

The upregulated genes of top significant pathways from each T-cell were subjected to PPI resulting in diverse networks of interacting proteins. The “CRC-Tumor vs CRC-Normal” group, including $CD8^+$ NKT-like cells exhibited a network of 132 proteins interacting with each other; memory $CD4^+$ T-cells showed a network of 135 interacting proteins; and naive $CD4^+$ T-cells exhibited a network of 22 interacting proteins. While, the “LM-Tumor vs LM-Normal” group, including memory $CD4^+$ T-cells exhibited a network of 40 interacting proteins; and naive $CD4^+$ T-cells showed a network of 34 interacting proteins. The protein networks of T-cells in “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups are shown in **Supplementary Figure S3 (A-E)**.

Subsequently, the protein networks of each T-cell were used to identify the most interacting genes (hub-genes) using the “Degree” method from Cytoscape. The protein networks from “CRC-Tumor vs CRC-Normal” group, including $CD8^+$ NKT-like cells showed top-5 hub-genes to be *NFKB1, CD44, CTLA4, HIF1A,* and *CXCL8*; exhibiting 36, 31, 28, 28, and 27 interactions in the protein network, respectively; memory $CD4^+$ T-cells exhibited *NFKB1, CXCL8, TGFB1, IL7R,* and *IL2RA*, showing 49, 34, 32, 31, and 30 interactions, respectively; lastly, naive $CD4^+$ T-cells exhibited *CCL5, IL7R, ANXA1, TIMP1,* and *LGALS3*. Among these *CCL5* and *IL7R* showed 10 and 8 interactions, respectively; while *ANXA1, TIMP1,* and *LGALS3* showed 6 interactions in the protein network.

Moreover, the “LM-Tumor vs LM-Normal” group, including memory $CD4^+$ T-cells exhibited *HSP90AA1, HSP90AB1, HSPA4, HSPA1A,* and *HSPA1B*, showing 29, 27, 25, 20, and 20 interactions, respectively; while naive $CD4^+$ T-cells showed *HSP90AA1, HSP90AB1, HSPA1A, HSPE1,* and *HSPD1*, exhibiting 24, 23, 14, 13, and 12 interactions, respectively. The hub-genes of T-cells in “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups are shown in **Figure 1 (K-O)**.

The aforementioned top-5 hub-genes of T-cells in “CRC-Tumor vs CRC-Normal” group indicated a chronic stimulation and potential exhaustion in $CD8^+$ NKT-like cells and memory $CD4^+$ T-cells in the primary tumor due to the overexpression of immune activation and inflammation-related genes; while naive $CD4^+$ T-cells exhibited migration and immune-regulatory genes, reflecting an early-stage activation;

whereas, cancer cells might be implicated in preventing complete activation of naive $CD4^+$ T-cells, keeping it in a non-effective state, promoting immune tolerance instead of clearance.

However, the top-5 hub-genes of T-cells in “LM-Tumor vs LM-Normal” group suggested an induction of stress adaptation and proteostasis pathways, promoting T-cell survival but suppressing activation, leading to a dysfunctional state. Interestingly, it was observed that *NFKB1* and *CXCL8* were common in $CD8^+$ NKT-like cells and memory $CD4^+$ T-cells; while *IL7R* was common among memory $CD4^+$ T-cells and naive $CD4^+$ T-cells of “CRC-Tumor vs CRC-Normal” group. In contrast, *HSP90AA1, HSP90AB1*, and *HSPA1A* genes were found common among memory $CD4^+$ T-cells and naive $CD4^+$ T-cells of the “LM-Tumor vs LM-Normal” group. This indicated the diverse roles of primary and metastatic cancer cells dysregulating different sets of genes in their respective niches.

3.5. Expression validation of T-cells hub-genes

The dataset, GSE136394 (metastatic state) used for the hub-genes expression validation consisted of major cell type lineages, including CD4T conv, CD8T, and T-prolif cells; while dataset GSE166555 (primary tumor) also consisted of CD4T conv, CD8T, and T-prolif cells, along with other cell types, including B cells, DC, Endothelial, Epithelial, Fibroblasts, Malignant, Mast, Mono/Macro, Myofibroblasts, and Plasma cells. Therefore, the expression levels of niche-specific hub-genes were observed in CD4T conv, CD8T, and T-prolif cells of respective datasets.

It was found that all the hub-genes of “CRC-Tumor vs CRC-Normal” group in GSE166555 dataset indicated an upregulation in tumor samples compared to normal samples, except for *CCL5* which showed lower expression (although upregulated) in tumor as compared to normal samples; hence *CCL5* was excluded from the analysis. Lastly, all the hub-genes from “LM-Tumor vs LM-Normal” group in GSE136394 dataset showed upregulation in the metastatic tumor samples; while for some hub-genes such as *HSP90AA1* and *HSP90AB1* showed upregulation in untreated tumor samples, while exhibited higher expression levels in treated tumor samples, indicating an immunotherapy response on these genes. The expression levels of T-cells hub-genes in heatmaps found in GSE136394 and GSE166555 datasets are shown in **Figure 2 (A-B)**.

Therefore, all the hub-genes from both groups (CRC-Tumor vs CRC-Normal and LM-Tumor vs LM-Normal), except for *CCL5*, were selected because of expression levels of hub-genes similar to T-cells in other datasets, specifically GSE136394 and GSE166555. The expression levels of selected hub-genes in this study for “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups are shown in **Figures 3 and 4**.

4. Discussion

There is a high rate of metastases in CRC patients, while approximately 20% of CRC patients have metastases at initial diagnosis (20). Recent studies have explored that the TIME of primary CRC and LM-

CRC is immunosuppressive, leading to the dysfunction of T-cells (21). Therefore, this study investigated the niche-specific (primary CRC and LM-CRC) genes of T-cells dysregulated due to cancer cells signaling within the TIME.

Subsequently, this study revealed top interacting hub-genes of $CD8^+$ NKT-like cells, memory $CD4^+$ T-cells, and naive $CD4^+$ T-cells in respective groups, including “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal”. The hub-genes of “CRC-Tumor vs CRC-Normal” group are reported to be implicated in various T-cells of tumors, such as *NFKB1* is reported to be majorly explored in Tregs of transgenic mice and its functional role has not been fully elucidated in human T-cells (22). However, it has been reported that activation of NF-κB pathway in cancer cells leads to the exhaustion of $CD8^+$ T-cells leading to LM of CRC (23).

Moreover, *CD44* has been reported to be expressed in $CD8^+$ T-cells and is implicated in migration and maintenance of cell shape during migration in a cancer model but not specifically in CRC; however, *CD44* is highly crucial in CRC cancer cells and play critical role in carcinogenesis, progression, metastasis, and resistance to therapy (24,25).

Tumor-infiltrating lymphocytes can express *CTLA4*, which leads to the attenuation of anti-tumoral immune responses and promoting metastases, whereas recent findings also demonstrate that *CTLA4* is also expressed on tumor cells shielding them from immune responses in CRC (26). *HIF1A* is reported to be involved in *HIF1A/ VEGF-A* response in T-cells, directly affecting cancer growth, progression, and vascularization in multiple cancer models, including the MC38 colon carcinoma model (27). *CXCL8* is not specifically reported in CRC; however, it is highly expressed in naive $CD4^+$ T-cells and is implicated in tumorigenesis and lymphoid metastasis (28).

A study revealed an immunosuppressive role of *TGFB1* by overexpressing PD-1 on T-cells, promoting tumor immune evasion (29). Furthermore, *IL7R* is reported to be highly expressed in memory T-cells and early activated $CD8^+$ T-cells, but not in terminally exhausted T-cells (30). Additionally, high expression of *IL2RA* (*CD25*), in Tregs, leads to suppressive functionality of Tregs and is associated with poor prognosis (31).

ANXA1 expression in T-cells is reported to promote tumor progression by enhancing the immunosuppressive T-cell responses, leading to suppression of anti-tumor immunity and enabling immune escape (32). A study suggested the presence of *TIMP1* in areas rich in immune activity in metastatic melanoma (33). *LGALS3* is overexpressed in Tregs of renal cell carcinoma (RCC) patients and is associated with the accumulation of immunosuppressive Tregs in tumor microenvironment, which may promote tumor immune evasion (34).

While, the hub-genes of “LM-Tumor vs LM-Normal” group are implicated in T-cells of various tumors, such as *HSP90AA1* and *HSP90AB1* are reported to be implicated in memory $CD4^+$ T-cells dysfunction in pancreatic ductal adenocarcinoma (PDAC) (35). *HSPA4* is not specifically reported in memory $CD4^+$ T-

cells; however, it is reported to be highly expressed in cancer cells of CRC, associated with poor prognosis, higher tumor stage, and increased metastatic potential (36). Subsequently, *HSPA1A* and *HSPA1B* are reported to be expressed in $CD8^+$ T-cells in primary and metastatic liver tumors (37). Lastly, *HSPE1* and *HSPD1* are reported to likely regulate apoptosis during T-cell activation (38).

The aforementioned implications of potential hub-genes in T-cells of CRC and various tumors reflect their functional role in suppressing immune cells, specifically T-cells within the TIME of primary CRC and LM-CRC. These overexpressed novel hub-genes may contribute to T-cell exhaustion, impaired activation, reduced cytotoxicity, and increased immunosuppressive signaling, which collectively promote tumor immune evasion and metastasis.

Additionally, previous studies of CRC have primarily focused on Tregs, cytotoxic $CD8^+$ T-cells, conventional helper $CD4^+$ T-cells, γδ T-cells, Th17 and engineering T-cells to treat tumors (39–41). However, there is a lack of molecular insights into T-cell subsets such as $CD8^+$ NKT-like cells, memory $CD4^+$ T-cells, and naive $CD4^+$ T-cells. This gap is particularly evident in primary and metastatic CRC, where their distinct roles and dysregulation within TIME remain largely unexplored.

Therefore, this study unraveled aforementioned novel T-cell specific biomarkers of primary CRC and LM-CRC specific niches, indicating potential targets for the development of effective immunotherapies, as the distinct immunosuppressive mechanisms operating in each tumor niche remain poorly defined.

However, this study was limited to a single dataset; larger cohorts are required to capture a broader diversity and higher number of T-cells to validate the findings more robustly. Additionally, further clinical validation is essential to confirm the functional and therapeutic relevance of the identified novel biomarkers.

Conclusion

This study identified T-cell specific ($CD8^+$ NKT-like cells, memory $CD4^+$ T-cells, and naive $CD4^+$ T-cells) novel biomarkers, including *NFKB1, CD44, CTLA4, HIF1A, CXCL8, TGFB1, IL7R, IL2RA, HSP90AA1, HSP90AB1, HSPA4, HSPA1A, HSPA1B, ANXA1, TIMP1, LGALS3, HSPE1,* and *HSPD1* in primary CRC and LM-CRC. These novel biomarkers might be implicated in the dysfunction of respective T-cells and promoting tumor immune evasion within TIME. Therefore, they might play a critical role as potential therapeutic targets for effective immunotherapies, which can be further validated through clinical investigations.

References

1. Gondal MN, Farooqi HMU. Single-Cell Transcriptomic Approaches for Decoding Non-Coding RNA Mechanisms in Colorectal Cancer. Non-Coding RNA. 2025 Apr;11(2):24.

2. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin. 2025;75(1):10–45.
3. Vatandoust S, Price TJ, Karapetis CS. Colorectal cancer: Metastases to a single organ. World J Gastroenterol. 2015 Nov 7;21(41):11767–76.
4. Jin Z, Li Y, Yi H, Wang M, Wang C, Du S, et al. Pathogenetic development, diagnosis and clinical therapeutic approaches for liver metastasis from colorectal cancer (Review). Int J Oncol. 2025 Mar 1;66(3):1–16.
5. Tsilimigras DI, Hyer MJ, Bagante F, Guglielmi A, Ruzzenente A, Alexandrescu S, et al. Resection of Colorectal Liver Metastasis: Prognostic Impact of Tumor Burden vsKRASMutational Status. J Am Coll Surg. 2021 Apr;232(4):590.
6. El Hage M, Su Z, Linnebacher M. Addressing Challenges in Targeted Therapy for Metastatic Colorectal Cancer. Cancers. 2025 Jan;17(7):1098.
7. Li Z, Deng L, Cheng M, Ye X, Yang N, Fan Z, et al. Emerging role of bile acids in colorectal liver metastasis: From molecular mechanism to clinical significance (Review). Int J Oncol. 2025 Mar 1;66(3):1–18.
8. Picard E, Verschoor CP, Ma GW, Pawelec G. Relationships Between Immune Landscapes, Genetic Subtypes and Responses to Immunotherapy in Colorectal Cancer. Front Immunol. 2020 Mar 6;11:369.
9. Sun J, Chen Y, Xu Z, Wang W, Li P. Notch signaling in the tumor immune microenvironment of colorectal cancer: mechanisms and therapeutic opportunities. J Transl Med. 2025 Mar 12;23(1):315.
10. Xiong Y, Wang Y, Tiruthani K. Tumor immune microenvironment and nano-immunotherapeutics in colorectal cancer. Nanomedicine Nanotechnol Biol Med. 2019 Oct 1;21:102034.
11. Zheng Z, Wieder T, Mauerer B, Schäfer L, Kesselring R, Braumüller H. T Cells in Colorectal Cancer: Unravelling the Function of Different T Cell Subsets in the Tumor Microenvironment. Int J Mol Sci. 2023 Jul 19;24(14):11673.
12. Tieng FYF, Lee LH, Ab Mutalib NS. Deciphering colorectal cancer immune microenvironment transcriptional landscape on single cell resolution – A role for immunotherapy. Front Immunol [Internet]. 2022 Aug 10 [cited 2025 Jun 3];13. Available from: https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2022.959705/full
13. Sayers EW, Beck J, Bolton EE, Brister JR, Chan J, Connor R, et al. Database resources of the National Center for Biotechnology Information in 2025. Nucleic Acids Res. 2024 Nov 11;53(D1):D20–9.
14. Nader K, Tasci M, Ianevski A, Erickson A, Verschuren EW, Aittokallio T, et al. ScType enables fast and accurate cell type identification from spatial transcriptomics data. Bioinformatics. 2024 Jun 27;40(7):btae426.
15. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics. 2015 May 22;16(1):169.
16. Fang Z, Liu X, Peltz G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. Bioinformatics. 2023 Jan 1;39(1):btac757.

17. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2022 Nov 12;51(D1):D638–46.
18. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. 2003 Nov;13(11):2498–504.
19. Han Y, Wang Y, Dong X, Sun D, Liu Z, Yue J, et al. TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. Nucleic Acids Res. 2022 Nov 2;51(D1):D1425–31.
20. Dong C, Zhang Y, Zeng J, Chong S, Liu Y, Bian Z, et al. FUT2 promotes colorectal cancer metastasis by reprogramming fatty acid metabolism via YAP/TAZ signaling and SREBP-1. Commun Biol. 2024 Oct 10;7(1):1–15.
21. Bell HN, Huber AK, Singhal R, Korimerla N, Rebernick RJ, Kumar R, et al. Microenvironmental ammonia enhances T cell exhaustion in colorectal cancer. Cell Metab. 2023 Jan 3;35(1):134-149.e6.
22. Sato Y, Osada E, Manome Y. Non-canonical NFKB signaling endows suppressive function through FOXP3-dependent regulatory T cell program. Heliyon [Internet]. 2023 Dec 1 [cited 2025 Jun 3];9(12). Available from: https://www.cell.com/heliyon/abstract/S2405-8440(23)10119-8
23. Rong D, Sun G, Zheng Z, Liu L, Chen X, Wu F, et al. MGP promotes CD8+ T cell exhaustion by activating the NF-κB pathway leading to liver metastasis of colorectal cancer. Int J Biol Sci. 2022 Mar 6;18(6):2345–61.
24. Ziranu P, Pretta A, Aimola V, Cau F, Mariani S, D'Agata AP, et al. CD44: A New Prognostic Marker in Colorectal Cancer? Cancers. 2024 Apr 19;16(8):1569.
25. Full article: Multifaceted regulation of T cells by CD44 [Internet]. [cited 2025 Jun 3]. Available from: https://www.tandfonline.com/doi/full/10.4161/cib.3.6.13495#d1e150
26. Derakhshani A, Hashemzadeh S, Asadzadeh Z, Shadbad MA, Rasibonab F, Safarpour H, et al. Cytotoxic T-Lymphocyte Antigen-4 in Colorectal Cancer: Another Therapeutic Side of Capecitabine. Cancers. 2021 May 17;13(10):2414.
27. Palazon A, Tyrakis PA, Macias D, Veliça P, Rundqvist H, Fitzpatrick S, et al. An HIF-1α/VEGF-A Axis in Cytotoxic T Cells Regulates Tumor Progression. Cancer Cell. 2017 Nov 13;32(5):669-683.e5.
28. Crespo J, Wu K, Li W, Kryczek I, Maj T, Vatan L, et al. Human naïve T cells express functional CXCL8 and promote tumorigenesis. J Immunol Baltim Md 1950. 2018 Jul 15;201(2):814–20.
29. Park BV, Freeman ZT, Ghasemzadeh A, Chattergoon MA, Rutebemberwa A, Steigner J, et al. TGF-β1-mediated Smad3 enhances PD-1 expression on antigen-specific T cells in cancer. Cancer Discov. 2016 Dec;6(12):1366–81.
30. Liang J, Zhu L, Li J, Wu K, Zhang M, Ma S, et al. Comprehensive analysis to identify IL7R as a immunotherapy biomarker from pan-cancer analysis to in vitro validation. Discov Oncol. 2024 Sep 30;15(1):509.

31. Bergsland CH, Jeanmougin M, Moosavi SH, Svindland A, Bruun J, Nesbakken A, et al. Spatial analysis and CD25-expression identify regulatory T cells as predictors of a poor prognosis in colorectal cancer. Mod Pathol. 2022 Sep 1;35(9):1236–46.
32. Gao K, Li X, Luo S, Zhao L. An overview of the regulatory role of annexin A1 in the tumor microenvironment and its prospective clinical application (Review). Int J Oncol. 2024 May 1;64(5):1–10.
33. Langguth M, Maranou E, Koskela SA, Elenius O, Kallionpää RE, Birkman EM, et al. TIMP-1 is an activator of MHC-I expression in myeloid dendritic cells with implications for tumor immunogenicity. Genes Immun. 2024 Jun;25(3):188–200.
34. Jeron A, Pfoertner S, Bruder D, Geffers R, Hammerer P, Hofmann R, et al. Frequency and Gene Expression Profile of Regulatory T Cells in Renal Cell Carcinoma. Tumor Biol. 2009 Jul 16;30(3):160–70.
35. Saleem MU, Sajid HA, Shabbir MI, Torres AR, Rai SK, Arshad MW. Novel Insights into T-Cell Exhaustion and Cancer Biomarkers in PDAC Using ScRNA-Seq [Internet]. bioRxiv; 2025 [cited 2025 Jun 4]. p. 2025.01.13.632574. Available from: https://www.biorxiv.org/content/10.1101/2025.01.13.632574v2
36. Zhang M, Dai W, Li Z, Tang L, Chen J, Chen C. HSPA4 Knockdown Retarded Progression and Development of Colorectal Cancer . Cancer Manag Res. 2021 Jun 14;13:4679–90.
37. Trehan R, Huang P, Zhu XB, Wang X, Soliman M, Strepay D, et al. SPP1 + macrophages cause exhaustion of tumor-specific T cells in liver metastases. Nat Commun. 2025 May 7;16(1):4242.
38. Wang M, Windgassen D, Papoutsakis ET. A global transcriptional view of apoptosis in human T-cell activation. BMC Med Genomics. 2008 Oct 23;1:53.
39. Rus Bakarurraini NAA, Kamarudin ,Ammar Akram, Jamal ,Rahman, and Abu N. Engineered T cells for Colorectal Cancer. Immunotherapy. 2024 Aug 21;16(14–15):987–98.
40. Zheng Z, Wieder T, Mauerer B, Schäfer L, Kesselring R, Braumüller H. T Cells in Colorectal Cancer: Unravelling the Function of Different T Cell Subsets in the Tumor Microenvironment. Int J Mol Sci. 2023 Jul 19;24(14):11673.
41. Olguín JE, Medina-Andrade I, Rodríguez T, Rodríguez-Sosa M, Terrazas LI. Relevance of Regulatory T Cells during Colorectal Cancer Development. Cancers. 2020 Jul 14;12(7):1888.

Figures

Cell Type Annotation, Volcano Plots of DEGs, Unique Genes, and Hub-Genes of T-cell Subsets in Primary and Metastatic CRC

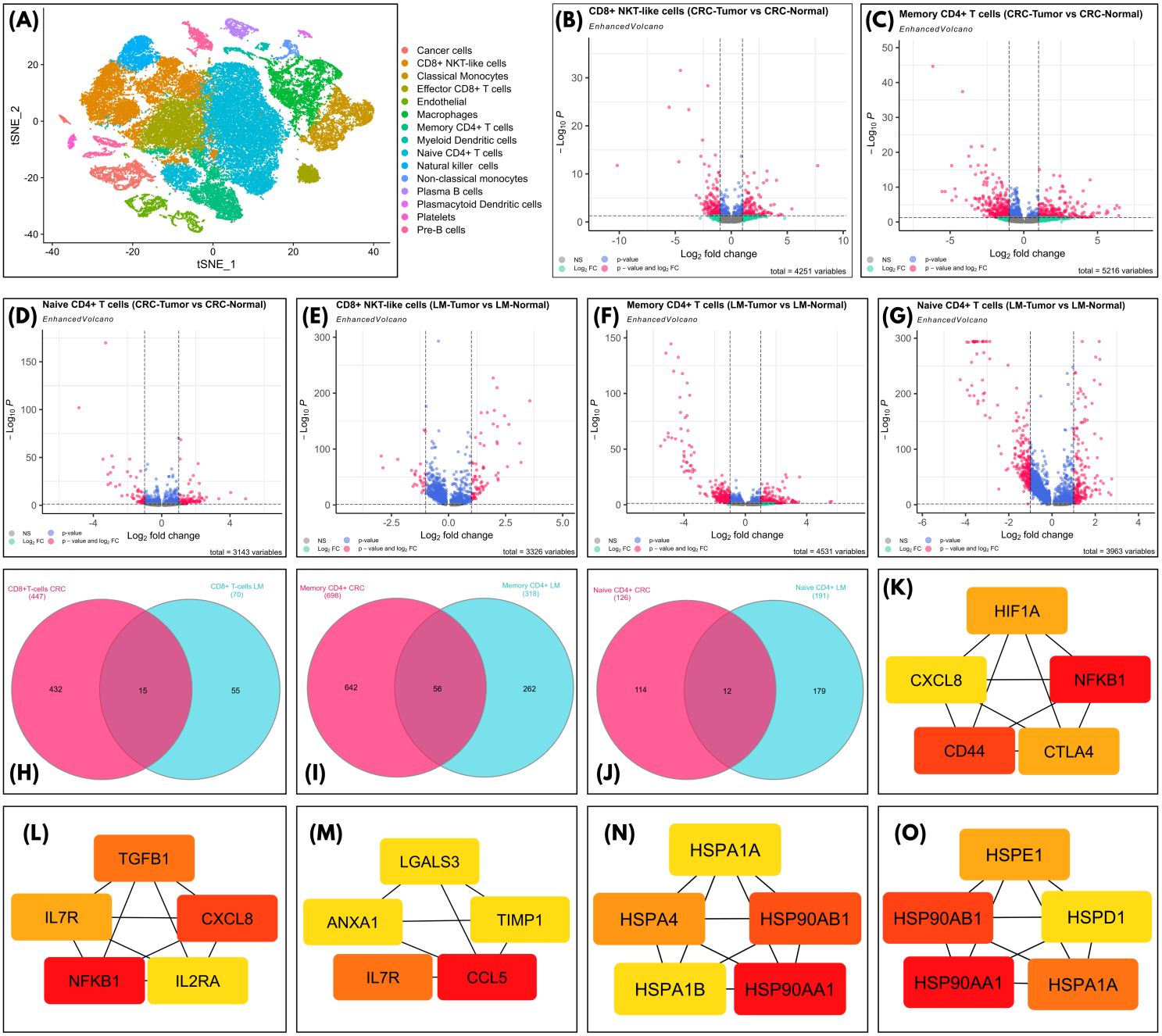


Figure 1

The cell type annotation of cell clusters, volcano plots of T-cell DEGs, common and unique genes of T-cells, and identification of the key hub-genes of T-cells; **(A)** tSNE plot of cell type annotations of CRC dataset showing total 15 cell types, including memory $CD4^+$ T cells, endothelial, pre-B cells, cancer cells, myeloid dendritic cells, plasma B cells, macrophages, naive $CD4^+$ T cells, plasmacytoid dendritic cells, classical monocytes, $CD8^+$ NKT-like cells, platelets, effector $CD8^+$ T cells, non-classical monocytes, and natural killer cells; **(B)** Volcano plot of $CD8^+$ NKT-like cells DEGs (CRC-Tumor vs CRC-Normal) showing a

total 4251 dysregulated (upregulated and downregulated) genes, showing significant p-value (0.05) and log2FC (±1) genes (dots) in red; significant p-value and log2FC < 1 and > -1 in blue; non-significant p-value and log2FC (±1) in green; and non-significant p-value and log2FC in grey; **(C)** Volcano plot of memory $CD4^+$ T-cells DEGs (CRC-Tumor vs CRC-Normal) showing a total 5216 dysregulated genes; **(D)** Volcano plot of naive $CD4^+$ T-cells DEGs (CRC-Tumor vs CRC-Normal) showing a total 3143 dysregulated genes; **(E)** Volcano plot of $CD8^+$ NKT-like cells DEGs (LM-Tumor vs LM-Normal) showing a total 3326 dysregulated genes; **(F)** Volcano plot of memory $CD4^+$ T-cells DEGs (LM-Tumor vs LM-Normal) showing a total 4531 dysregulated genes; **(G)** Volcano plot of naive $CD4^+$ T-cells DEGs (LM-Tumor vs LM-Normal) showing a total 3963 dysregulated genes; **(H)** Venn diagram showing common and unique genes of $CD8^+$ NKT-like cells among “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups; **(I)** Venn diagram showing common and unique genes of memory $CD4^+$ T-cells among “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups; **(J)** Venn diagram showing common and unique genes of naive $CD4^+$ T-cells among “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups; **(K)** Top-5 hub-genes of $CD8^+$ NKT-like cells (CRC-Tumor vs CRC-Normal), including *NFKB1, CD44, CTLA4, HIF1A*, and *CXCL8*; **(L)** Top-5 hub-genes of memory $CD4^+$ T-cells (CRC-Tumor vs CRC-Normal), including *NFKB1, CXCL8, TGFB1, IL7R*, and *IL2RA*; **(M)** Top-5 hub-genes of naive $CD4^+$ T-cells (CRC-Tumor vs CRC-Normal), including *CCL5, IL7R, ANXA1, TIMP1*, and *LGALS3*; **(N)** Top-5 hub-genes of memory $CD4^+$ T-cells (LM-Tumor vs LM-Normal), including *HSP90AA1, HSP90AB1, HSPA4, HSPA1A,* and *HSPA1B*; **(O)** Top-5 hub-genes of naive $CD4^+$ T-cells (LM-Tumor vs LM-Normal), including *HSP90AA1, HSP90AB1, HSPA1A, HSPE1,* and *HSPD1*

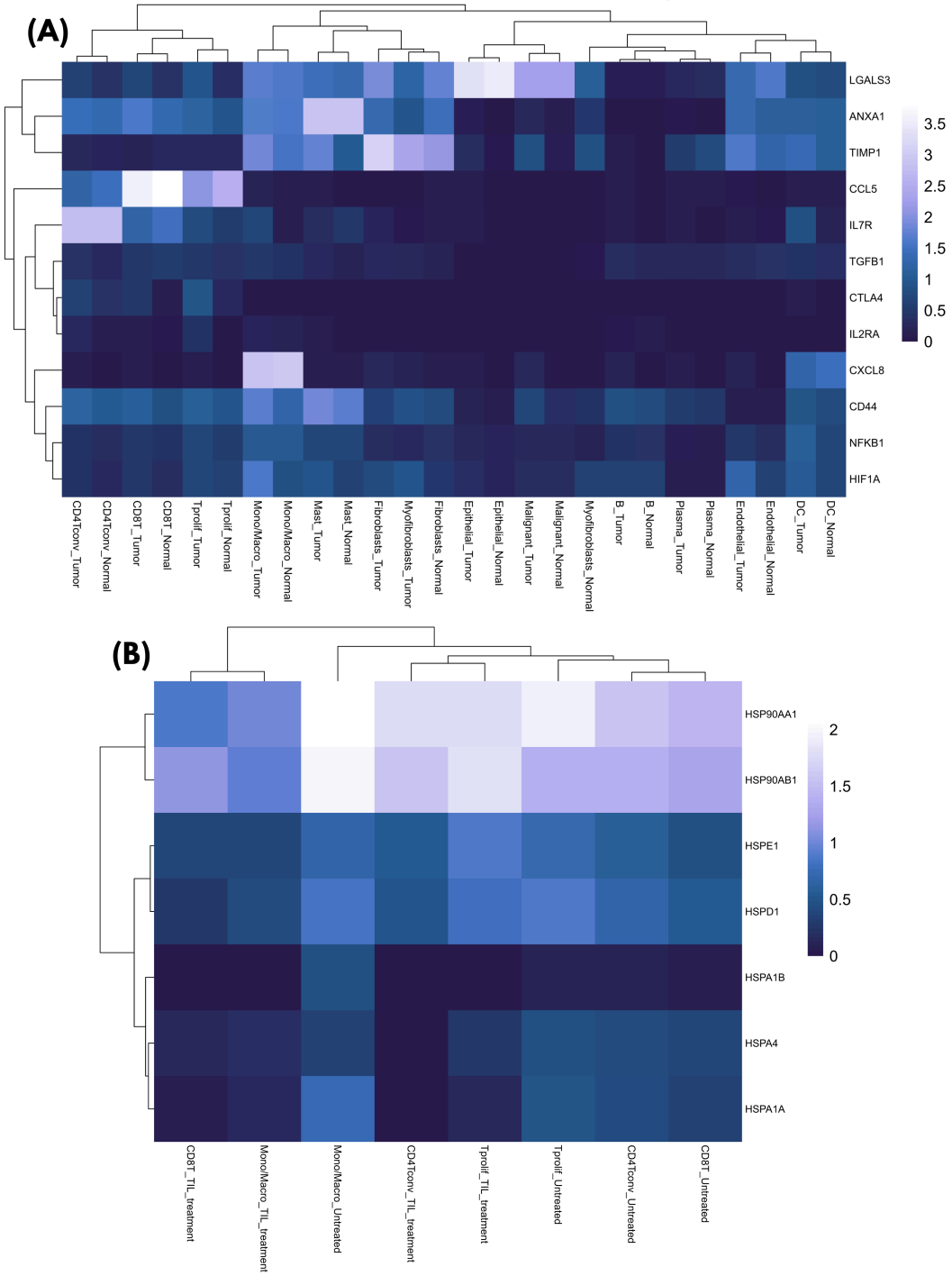


Figure 2

The expression validation of “CRC-Tumor vs CRC-Normal” and “LM-Tumor vs LM-Normal” groups top-5 hub-genes using TISCH2; **(A)** Heatmap of “CRC-Tumor vs CRC-Normal” group T-cells hub-genes exhibiting expression levels in tumor and normal samples of CD8T, CD4T conv, and T-prolif cells along with other cell types from GSE166555 dataset, showing all the hub-genes expression levels to be elevated in tumor samples compared to normal samples of T-cells except for *CCL5* which exhibited higher expression in normal samples compared to tumor samples; **(B)**Heatmap of “LM-Tumor vs LM-

Normal” group T-cells hub-genes exhibiting expression levels in treated and untreated metastatic tumor samples of CD8T, CD4T conv, and T-prolif cells from GSE136394 dataset, showing all the hub-genes expression levels to be elevated in untreated metastatic tumor samples, while some hub-genes such as *HSP90AA1* and *HSP90AB1* expressed higher expression levels in treated metastatic tumor samples, indicating a response of immunotherapy on these genes

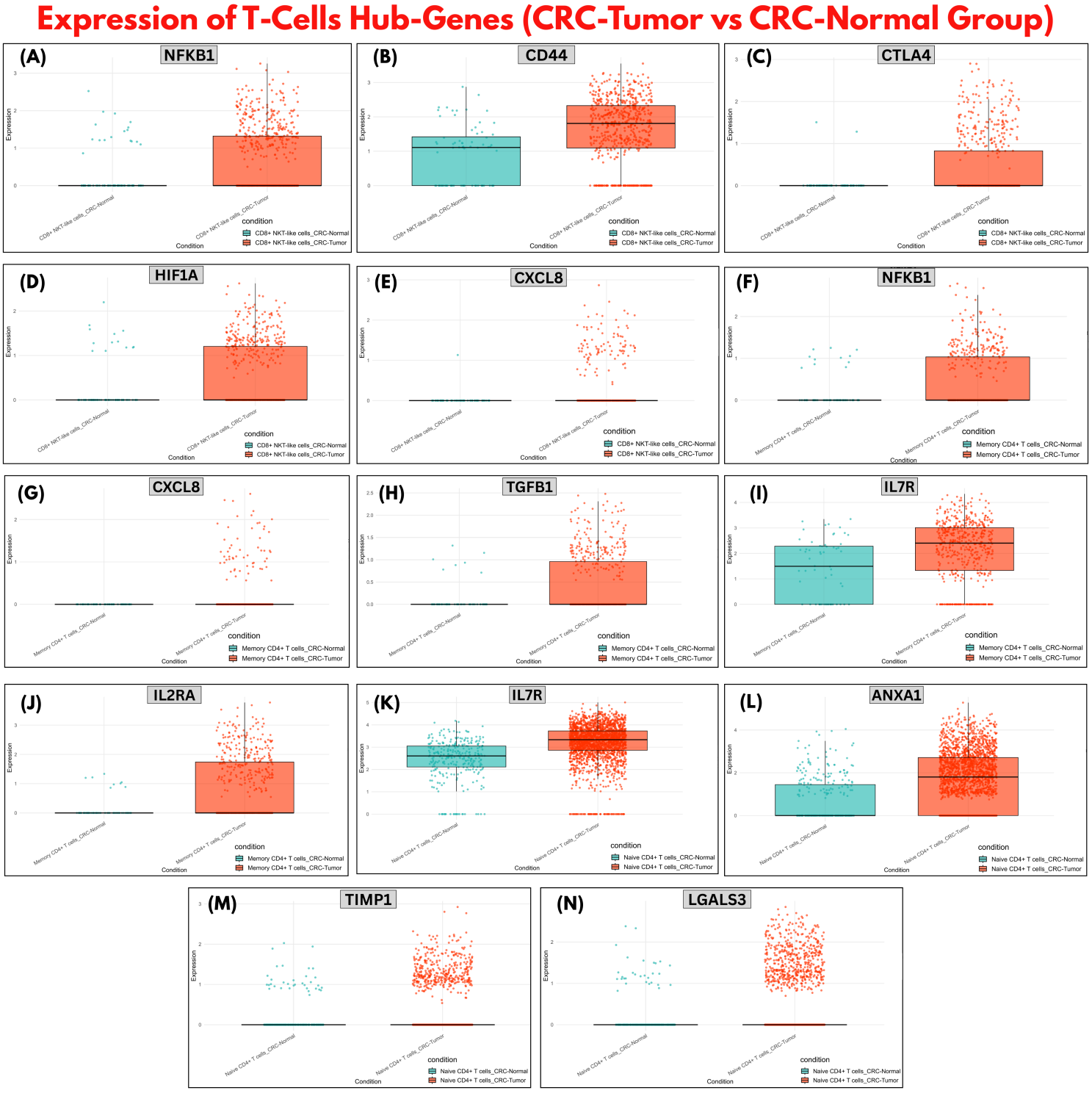


Figure 3

Expression levels of selected hub-genes of $CD8^+$ NKT-like cells, memory $CD4^+$ T-cells, and naive $CD4^+$ T-cells in “CRC-Tumor vs CRC-Normal” group; **(A-E)** Expression levels of *NFKB1, CD44, CTLA4, HIF1A,* and *CXCL8* in $CD8^+$ NKT-like cells, showing high expression in CRC-Tumor condition compared to CRC-Normal; **(F-J)** Expression levels of *NFKB1, CXCL8, TGFB1, ILR,* and *IL2RA* in memory $CD4^+$ T-cells, showing high expression in CRC-Tumor condition compared to CRC-Normal; **(K-N)** Expression levels of *IL7R, ANXA1, TIMP1,* and *LGALS3* in naive $CD4^+$ T-cells, showing high expression in CRC-Tumor condition compared to CRC-Normal

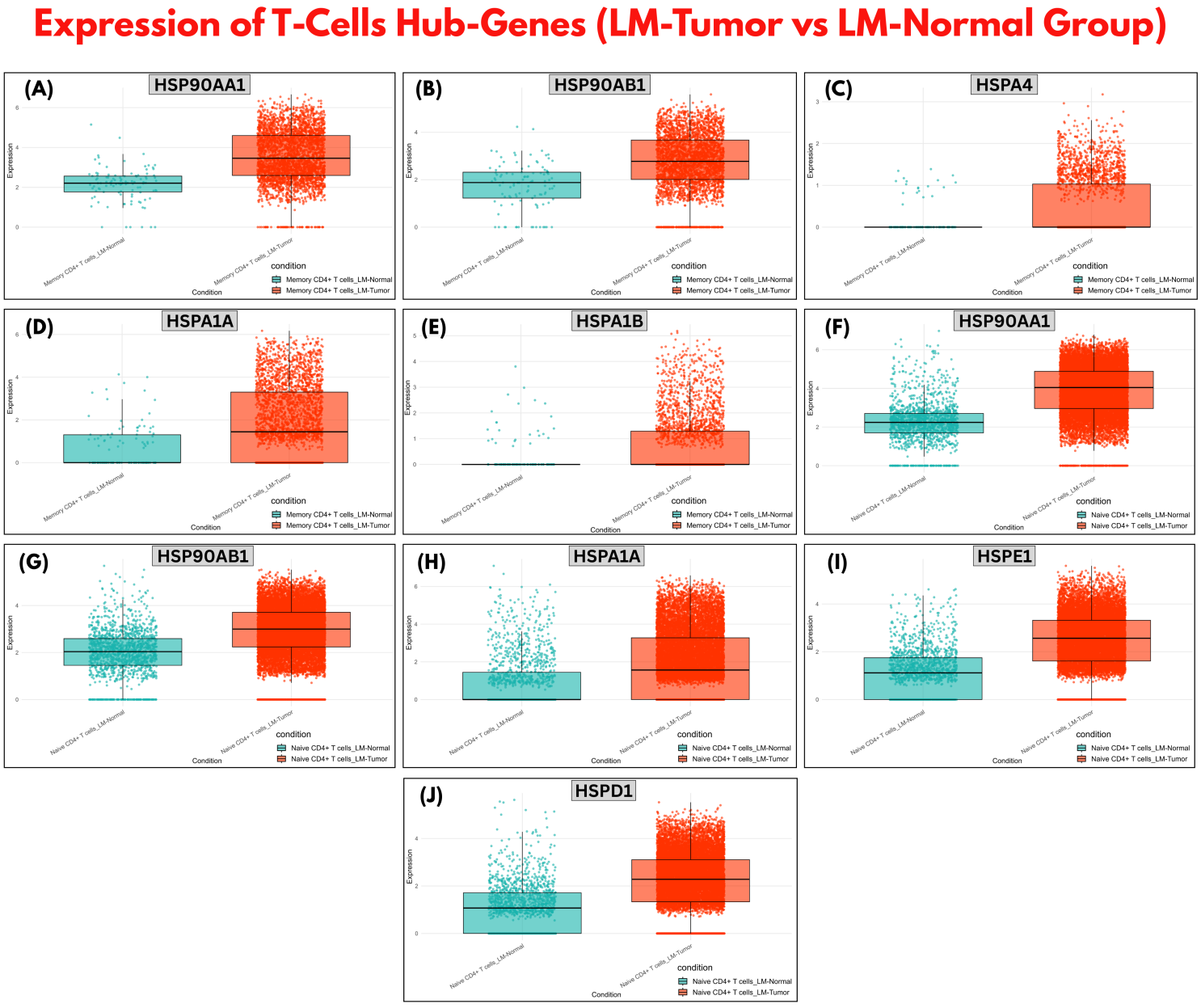


Figure 4

Expression levels of selected hub-genes of memory $CD4^+$ T-cells and naive $CD4^+$ T-cells in “LM-Tumor vs LM-Normal” group; **(A-E)** Expression levels of *HSP90AA1, HSP90AB1, HSPA4, HSPA1A,* and *HSPA1B* in memory $CD4^+$ T-cells, showing high expression in LM-Tumor compared to LM-Normal condition; **(F-J)** Expression levels of *HSP90AA1, HSP90AB1, HSPA1A, HSPE1* and *HSPD1* in naive $CD4^+$ T-cells, showing high expression in LM-Tumor compared to LM-Normal condition

Supplementary Files

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

- GA.png
- SupplementalDocument.docx
- S1Fig.png
- S2Fig.png
- S3Fig.png