scientific reports

Check for updates

OPEN

Transcriptome combined with single-cell sequencing explored prognostic markers associated with T cell exhaustion characteristics in head and neck squamous carcinoma

Jie Liu[1], Penghui Li[2], Yuanyuan Zhang[1] & Lian Zheng[1,3]✉

Head and neck squamous cell carcinoma (HNSC) ranks among the most prevalent cancers worldwide, characterized by significant heterogeneity and a complex immune microenvironment. T cell exhaustion is pivotal in the pathogenesis of HNSC, where depleted T cells exhibit reduced proliferative capacity and diminished effector function, facilitating tumor immune escape and subsequent disease progression. A thorough understanding of the primary mechanisms driving T cell depletion within the tumor microenvironment is essential for enhancing the efficacy of immunotherapeutic approaches in HNSC, with profound implications for patient outcomes. In this study, a single-cell atlas of HNSC was constructed, enabling an in-depth analysis of T cell heterogeneity. The differentiation trajectory of T cells, transitioning from normal tissue to HNSC, was characterized, revealing a predisposition toward depletion in the C2 T cell subgroup. A subsequent cross-analysis of significantly upregulated differentially expressed genes in the C2 T cell subset identified five characteristics pertinent to T cell C2, which informed the development of a clinical prognostic model. Additionally, maximum half inhibitory concentration (IC_{50}) values for various pharmacological agents were calculated, leading to the identification of eleven drugs relevant to the risk model, providing an intriguing starting point for further work in personalized cancer treatment. However, certain limitations of this study must be acknowledged. While T cell heterogeneity and differentiation trajectories were mapped, the interrelationships among these subpopulations remain poorly understood. Further research is required to elucidate the specific biological processes and molecular evolutionary mechanisms involved. The insights from this study provide a valuable foundation for future investigations into the molecular mechanisms and immune landscape associated with the progression from normal tissue to malignant squamous cell carcinoma.

Keywords Head and neck squamous cell carcinoma, T cell exhaustion, Transcriptome, Single-cell sequencing, Prognosis

Head and neck squamous cell carcinoma (HNSC) encompasses a diverse spectrum of solid tumors that arise within the upper respiratory and digestive systems, marked by significant genetic complexity and aggressive behavior[1,2]. Key risk factors, such as tobacco use and heavy alcohol consumption, are strongly associated with HNSC[3]. This group of tumors is linked to high morbidity and mortality rates, primarily due to limited therapeutic efficacy, drug resistance, and a high incidence of recurrence and metastasis, all of which contribute to an unfavorable prognosis[4,5]. HNSC is often diagnosed at an advanced stage, characterized by locally advanced

[1]Department of Oral and Maxillofacial Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, Henan, China. [2]Department of Gastrointestinal surgery, The First Affiliated Hospital, College of Clinical Medicine, Henan University of Science and Technology, Luoyang 471000, Henan, China. [3]The First Affiliated Hospital of Zhengzhou University, NO.1 Jianshedong Road, Zhengzhou, Henan 450052, China. ✉email: zhenglian0726@163.com

disease or distant metastasis[6]. Immune cell infiltrates within the tumor microenvironment play a pivotal role in tumor progression and significantly influence clinical outcomes in patients with cancers[7]. Immunotherapy, recognized for its low toxicity, high specificity, and targeted mechanisms, effectively leverages the host immune system to inhibit tumor evasion[8,9]. It can be administered either as monotherapy or in combination with standard therapeutic approaches and is considered one of the effective treatments for advanced HNSC[10,11].

The immune system plays a critical role in monitoring the initiation, progression, and advancement of HNSC[12]. While immunotherapy has shown potential in prolonging survival for some patients with HNSC, response rates remain suboptimal[13]. Furthermore, HNSC exhibits various mechanisms that facilitate immune evasion[14,15]. T cell exhaustion is a common phenomenon in cancer, characterized by a progressive decline in T cell functionality and an increased expression of inhibitory receptors, such as CTLA-4[16]. Depleted T cells exhibit reduced proliferative capacity and diminished effector function, impairing their ability to recognize and eliminate tumor cells[17]. This leads to immune evasion by tumors and further disease progression[18]. However, Wu et al. demonstrated that blocking T-cell immunoglobulin and ITIM domain (TIGIT)/CD155 signaling can reverse T cell exhaustion in HNSC and enhance anti-tumor activity[19]. Therefore, elucidating the primary mechanisms driving T cell exhaustion within the tumor microenvironment could enhance the effectiveness of immunotherapy for HNSC and is of significant clinical relevance.

In this study, a comprehensive single-cell atlas of HNSC was developed by integrating single-cell and transcriptomic sequencing data from HNSC samples and adjacent non-tumor tissues. T cell heterogeneity was also analyzed to delineate the differentiation trajectories of T cells transitioning from normal tissue to HNSC. Additionally, we examined the differences in T cell communication between HNSC and normal tissue samples. An in-depth analysis was conducted on the characteristics of T cell depletion observed in HNSC, and the implications of these characteristics on patient prognosis were explored. The findings of this study could facilitate the identification of novel prognostic biomarkers and therapeutic targets, thus laying the groundwork for advancing personalized immunotherapy strategies.

Results

HNSC cell single cell atlas

scRNA-seq samples from 19 HNSC cases, available in the GEO database under accession number GSE181919, were initially obtained. After undergoing preprocessing steps, including cell filtration, normalization, dimensionality reduction, and clustering, a total of 23,509 cells were identified and categorized into 15 distinct clusters. To improve classification accuracy, established marker genes were reclassified into nine primary cell types based on their expression profiles within these clusters. The identified cell types included T cells, fibroblasts, macrophages, malignant cells, endothelial cells, B/plasma cells, dendritic cells, mast cells, and myocytes (Fig. 1A-C). A comparative analysis was conducted to assess the distribution of these cells across 13 primary cancer tissue samples (CA) and 6 non-peritumoral normal tissue samples (NL), with their relative proportions illustrated in Fig. 1D. The scale map analysis revealed that T cells constitute a significant portion of the cellular composition within the examined samples.

T cell heterogeneity

T cells play a critical role in the immune system by facilitating the activation and inhibition of other immune cells, directly engaging in target cell destruction, and modulating cytokine production[20,21]. A subsequent UMAP clustering analysis identified six distinct T cell subgroups: C1, C2, C3, C4, C5 and C6 (Fig. 2A). To further explore the role of these T cell subsets in tumor progression, the Monocle2 algorithm was utilized to construct differentiation trajectories, followed by a pseudotime analysis. NL samples were used as the initial reference point, with CA samples representing the final stage of progression (Fig. 2B and Figure S1). The analysis revealed two predominant T cell trajectories during the transition from non-tumor surrounding normal tissue to HNSC (Fig. 2C). Pseudo-time trajectory analysis revealed T cell exhaustion occurring as pseudo-time progressed. Throughout this process, immune checkpoints associated with T cell exhaustion, such as CTLA4, lymphocyte activating gene 3 (LAG3), and TIGIT, progressively increased (Fig. 2D). These checkpoints were predominantly found within the T cell C2 cluster, indicating that ongoing stimulation during tumor progression led to T cell exhaustion. These exhausted T cells lose their ability to effectively recognize and destroy tumor cells, potentially allowing the tumor to evade immune responses and continue growing[22]. The FindAllMarkers function was employed to identify genes specifically expressed in different T cell subsets, with parameters set to logfc.threshold = 0.25, min.pct = 0.25, and only.pos = T. This analysis identified 337 marker genes for T cell C2, which were subsequently used to investigate the relationship between T cell depletion characteristics and head and neck squamous cancer prognosis. Figure S2 presents a heatmap of the Top20 differentially expressed genes for each T cell subpopulation. Refer to "metadata.csv" for a specific metadata information on T cell populations and refer to "cluster_Tcell_celltype_DEG.txt" for a specific genes characteristic of T cell subpopulations.

Cell-cell interactions between T cells

In multicellular organisms, cellular activities are fundamentally governed by interactions between cells, primarily mediated by multi-subunit protein complexes. CellChat is a computational tool designed to infer and analyze communication networks between cells using single-cell transcriptome data, all within a framework that is accessible and easy to interpret[23]. In this study, CellChat packets were employed to examine variations in T cell communication between samples derived from HNSC and normal tissues. A notable increase in the number of intercellular interactions was observed in HNSC cells compared to normal cells (Fig. 3A). Additionally, an analysis of the overall information flow within the network between HNSC and normal samples revealed significant differences, allowing for the ranking of key signaling pathways. The orange apical signaling pathway

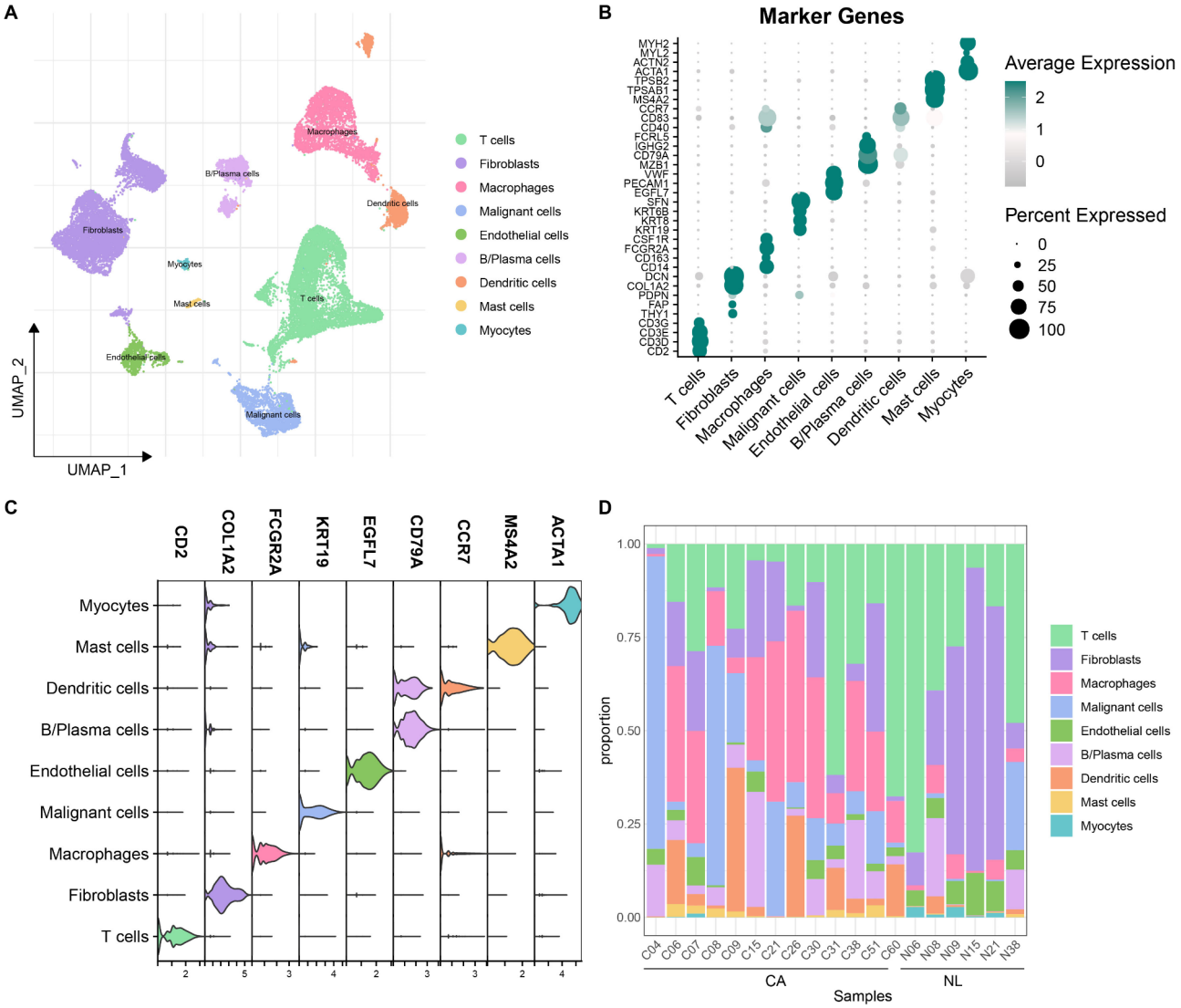

Fig. 1. Single-cell atlas of head and neck squamous cell carcinoma. (**A**) UMAP of annotated cell subsets. (**B**) High expression genes specific to different cell types. (**C**) Violin plot displaying marker gene expression levels across cell types. (**D**) Proportion of cell subpopulations in various samples.

was predominantly enriched in HNSC cells, while the green signaling pathway was more enriched in normal cells (Fig. 3A). T cell C2 displayed substantial intercellular communication within the context of HNSC and received regulatory signals from other cellular entities (Fig. 3B). A bubble diagram illustrating the interactions of T cell C2 with other immune cells is presented in Fig. 3C, while Fig. 3D shows the influence of various immune cells on T cell C2. Comparative analysis of ligand-receptor interactions in HNSC versus normal samples revealed that the LGALS9-CD45 interaction was exclusively present in the tumor microenvironment. Both LGALS9 and CD45 play roles in modulating T cell activity and regulating immunosuppressive mechanisms. Notably, elevated expression levels of LGALS9 were observed in tumor cells and tumor-associated immune cells, which impair the functionality of effector T cells and facilitate tumor immune evasion through interaction with CD45 on the surface of T cells[24,25].

Identification of T cell C2-related genes through WGCNA

Marker genes with elevated expression specific to T cell C2 were utilized as a reference gene set in this study. The ssGSEA methodology was then applied to calculate T cell C2-related scores within the TCGA-HNSC dataset, revealing that these scores were significantly higher in cancer samples compared to normal samples (Fig. 4A). The R software package WGCNA was used to identify modules associated with T cell C2 in patients with HNSC utilizing the TCGA-HNSC dataset. The pickSoftThreshold function was employed to establish a soft threshold of $\beta = 16$, leading to the identification of seven distinct modules (Fig. 4B). Correlation scores derived from ssGSEA and clinical outcomes were used as traits for dataset analysis, resulting in the generation of a heatmap illustrating the correlation among various modules. The green module, which exhibited the strongest correlation, was

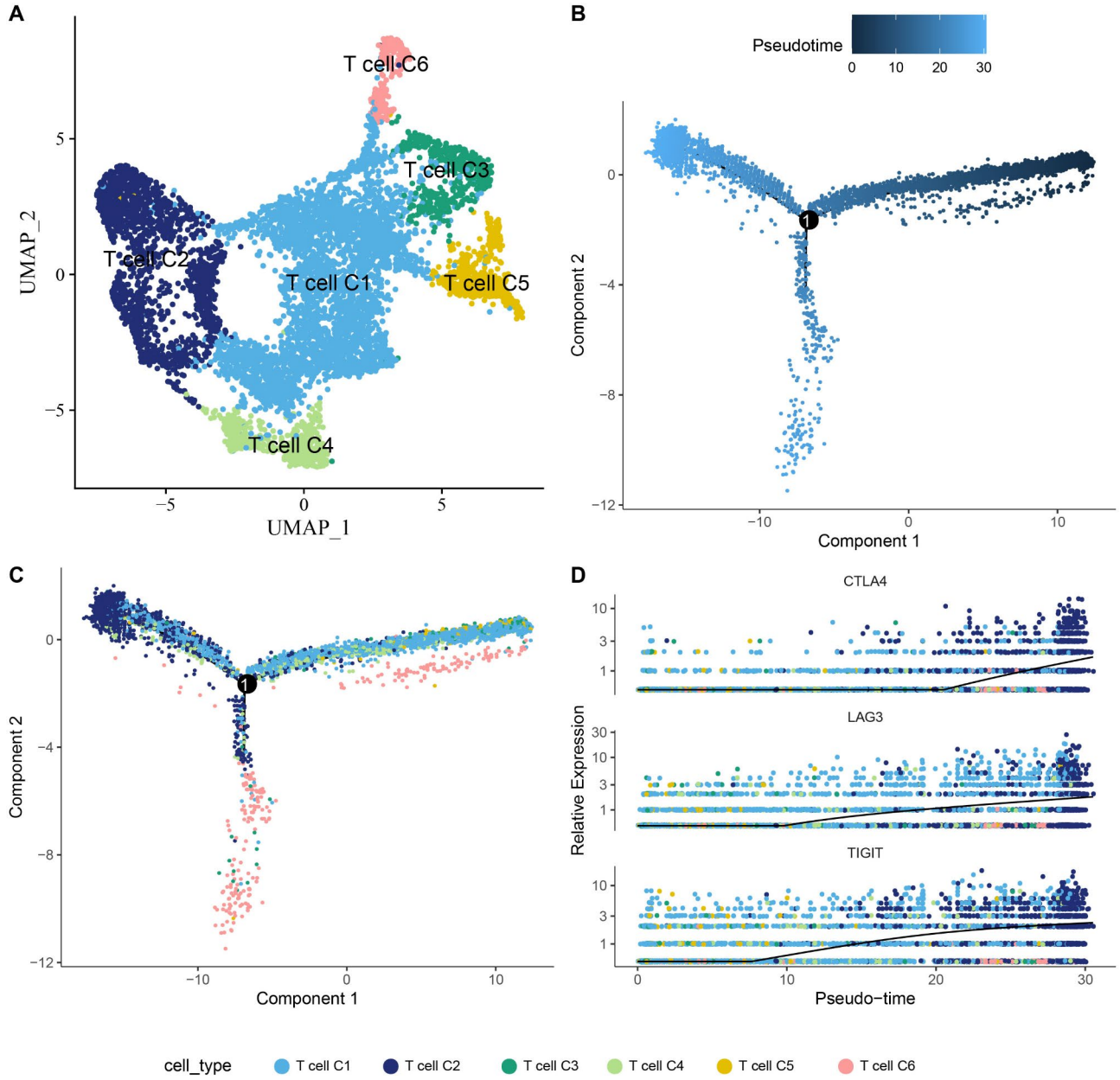

Fig. 2. Heterogeneity and differentiation trajectories of T cells. (**A**) UMAP dimensionality reduction following T cell clusters annotation. (**B**) Differentiation trajectory of T cells from normal tissue to HNSC, colored by pseudotime. (**C**) Differentiation trajectory of T cells from normal tissue to HNSC, colored by T cell clusters. (**D**) Heatmap of gene expression levels associated with T cell exhaustion.

identified as a specific module containing 558 genes (Fig. 4C). Among these, 477 genes (with GS ≥ 0.3 and |MM| ≥ 0.6) were classified as module genes highly associated with the T cell C2 trait (Fig. 4D).

Establishment and validation of clinical prognostic model

The limma analysis was applied to identify differentially expressed genes between Tumor and Normal samples within the TCGA-HNSC cohort, using a threshold of FDR < 0.05 and |logFC| > log2(1.5). The 5903 differentially expressed genes in HNSC were intersected with the gene set of T cell C2 highly related modules identified by WGCNA analysis to obtain 248 candidate genes for further analysis.

Subsequently, the 490 cancer samples from TCGA-HNSC were divided into a training set and a test set in an 8:2 ratio for model construction. The clinical characteristics of both sets are detailed in Supplementary Table 1. A Chi-square test comparing the clinical features of the two groups revealed no significant differences ($p > 0.05$), confirming that the grouping was random and appropriate.

Univariate COX regression analysis was then performed on the 248 candidate genes within the training set, identifying 28 genes with significant prognostic relevance ($p < 0.05$). To refine the selection, lasso regression was

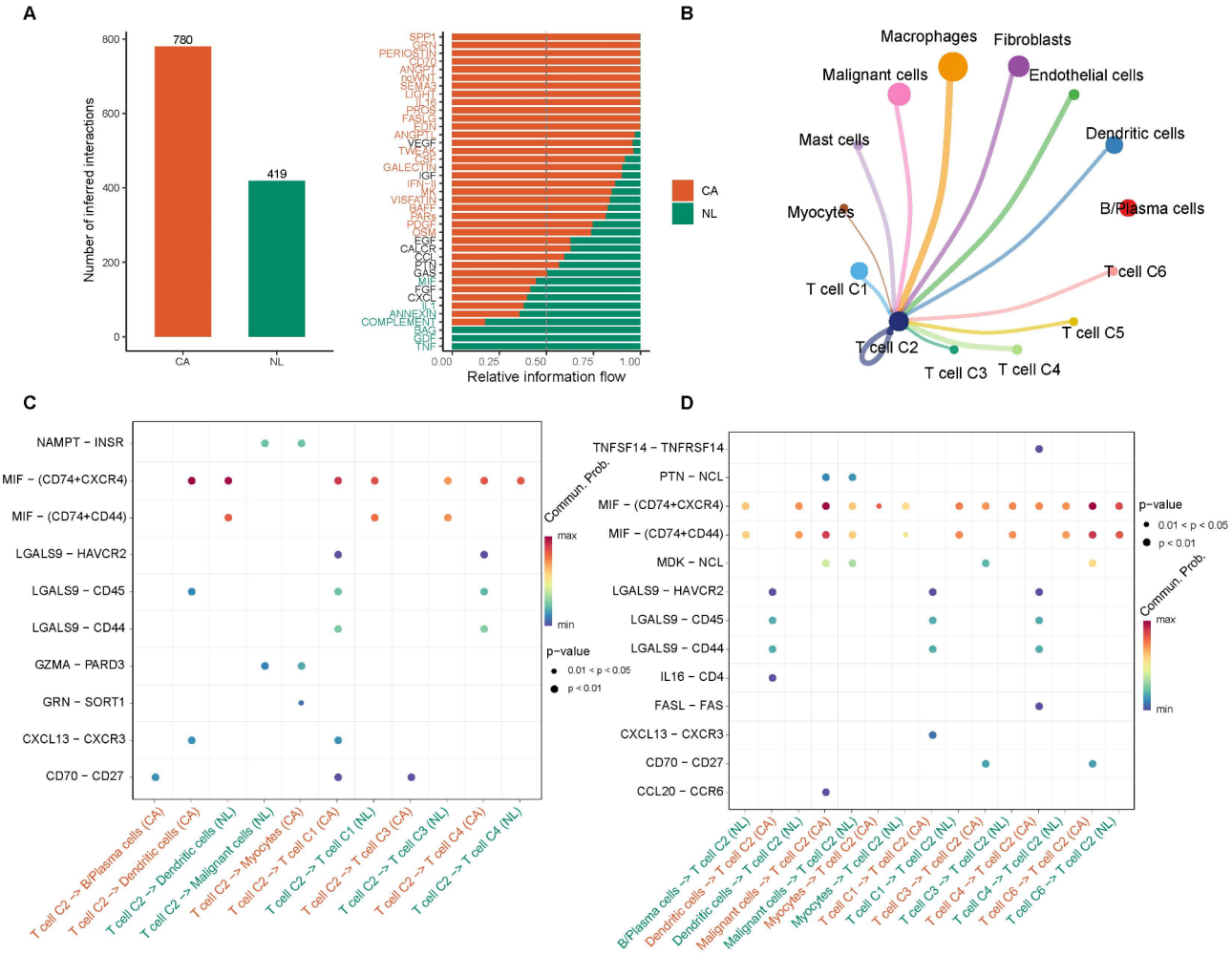

Fig. 3. Intercellular communication of T cells. (**A**) Comparative analysis of the number of cell communication interactions and pathway signal intensity between the CA and NL groups. (**B**) Communication diagram of T cell C2 as a signal receiver. (**C**) Bubble diagram illustrating T cell C2 interactions with immune cells in the CA and NL groups. (**D**) Bubble diagram showing the interaction of immune cells with T cell C2 in the CA and NL groups.

conducted using the R package 'glmnet,' further narrowing down the significant prognostic genes. Lasso method is a kind of compression estimation. It produces a more refined model by constructing a penalty function that compresses some coefficients and sets some coefficients to zero[26]. This method retains the advantage of subset shrinkage, is a biased estimation of complex collinearity data, and can realize the selection of variables while parameter estimation, and better solve the multicollinearity problem in regression analysis[27]. The trajectory of change for each independent variable was analyzed (Fig. 5A-B), showing that as lambda increased, the number of independent variable coefficients approaching zero also increased. A 10-fold cross-validation approach was employed to construct the model and determine the confidence interval for each lambda value[28]. The model's optimal performance was achieved at lambda = 0.0172, leading to the selection of 12 genes at lambda = 0.0186 for further analysis. Building on the lasso regression results, stepwise multivariate regression analysis identified five signature genes, which served as preliminary indicators for further evaluation (Fig. 5C). The final model formula is as follows: RiskScore = 0.385*FCGR2A-0.249*ICOS-0.317*SH2D3C + 0.344*SLC29A3-0.25*TNFRSF4.

Using the formula established by our risk model, the RiskScore for each sample in the training set was calculated. The RiskScore underwent z-score normalization, allowing for the categorization of patients into high- and low-risk groups based on a threshold of 0. Comparative analysis of overall survival rates revealed that patients in the high RiskScore group had a significantly lower survival rate compared to those in the low RiskScore group, with a marked difference in overall survival time ($p < 0.05$). Furthermore, the timeROC package in R was utilized to perform the Receiver Operating Characteristic (ROC) analysis of the RiskScore, assessing the model's predictive efficiency for survival over one to five years (Fig. 5D). The results demonstrated a high Area Under the Curve (AUC) value, indicating strong model performance. To validate the robustness of our clinical prognostic model, we conducted tests in both the test set and the GSE41613 cohort, with results consistent with those observed in the training set (Fig. 5E-F). Refer to "tcga.xlsx" for genes of significant prognostic relevance.

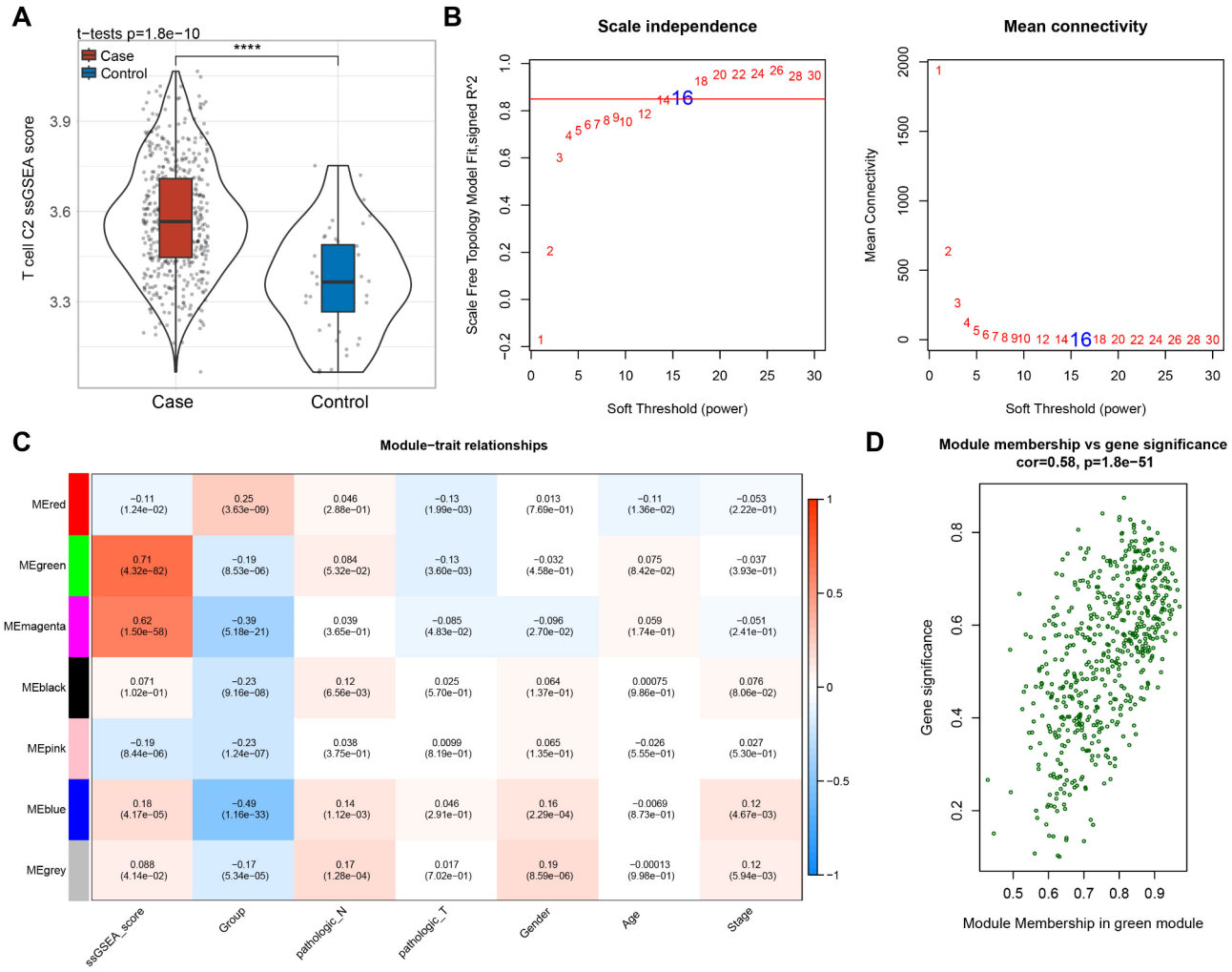

Fig. 4. Identification of T cell C2-related genes through WGCNA. (**A**) Box plots of ssGSEA results showing significantly higher T cell C2-related scores in cancer samples compared to normal samples. (**B**) WGCNA soft threshold filtering diagram, with 16 chosen as the soft threshold for constructing a topological network. (**C**) Module-trait correlation chart, where the horizontal axis represents traits, the vertical axis represents modules, numbers in boxes indicate correlation coefficients, and numbers in parentheses denote significance P-values, with red indicating positive correlation and blue indicating negative correlation. (**D**) GS-MM diagram of the green module, where the x-axis is MM (module membership), representing the correlation between each gene and the characteristic gene; the y-axis is GS (gene significance), representing the correlation between genes and traits in the module.

Risk models with immune microenvironment and chemotherapy drug sensitivity

To elucidate the relationship between RiskScore and the immune microenvironment in patients with HNSC, MCP-counter software was utilized to assess various immune cell populations. The analysis revealed significant enrichment of distinct immune and stromal cell types in both high-risk and low-risk patient cohorts (Fig. 6A). Additionally, the maximum IC_{50} values for a range of pharmacological agents within the TCGA-HNSC dataset were calculated to identify drugs with significant differential responses, facilitating a deeper exploration of the relationship between the prognostic model and drug sensitivity. The findings indicated that 11 drugs exhibited a significant correlation with RiskScore (FDR < 0.05 & |cor| > 0.3). The correlation analysis results of RiskScore and expression of key genes in the model with IC50 are shown in Fig. 6. We found that in the high RiskScore group, the IC_{50} value for PF.4708671_1129 was elevated. This suggests that patients with a high RiskScore may have reduced sensitivity to this drug, whereas those in the low-risk group showed increased sensitivity to this class of pharmacological agents (Fig. 6C-D). PF.4708671_1129, serving as an S6K1 inhibitor, plays a crucial role in investigating the impact of the mTOR/S6K1 signaling pathway on cell growth, metabolism, and tumorigenesis[29,30]. Additionally, it holds significant experimental value in exploring mechanisms of tumor therapy and metabolic diseases.

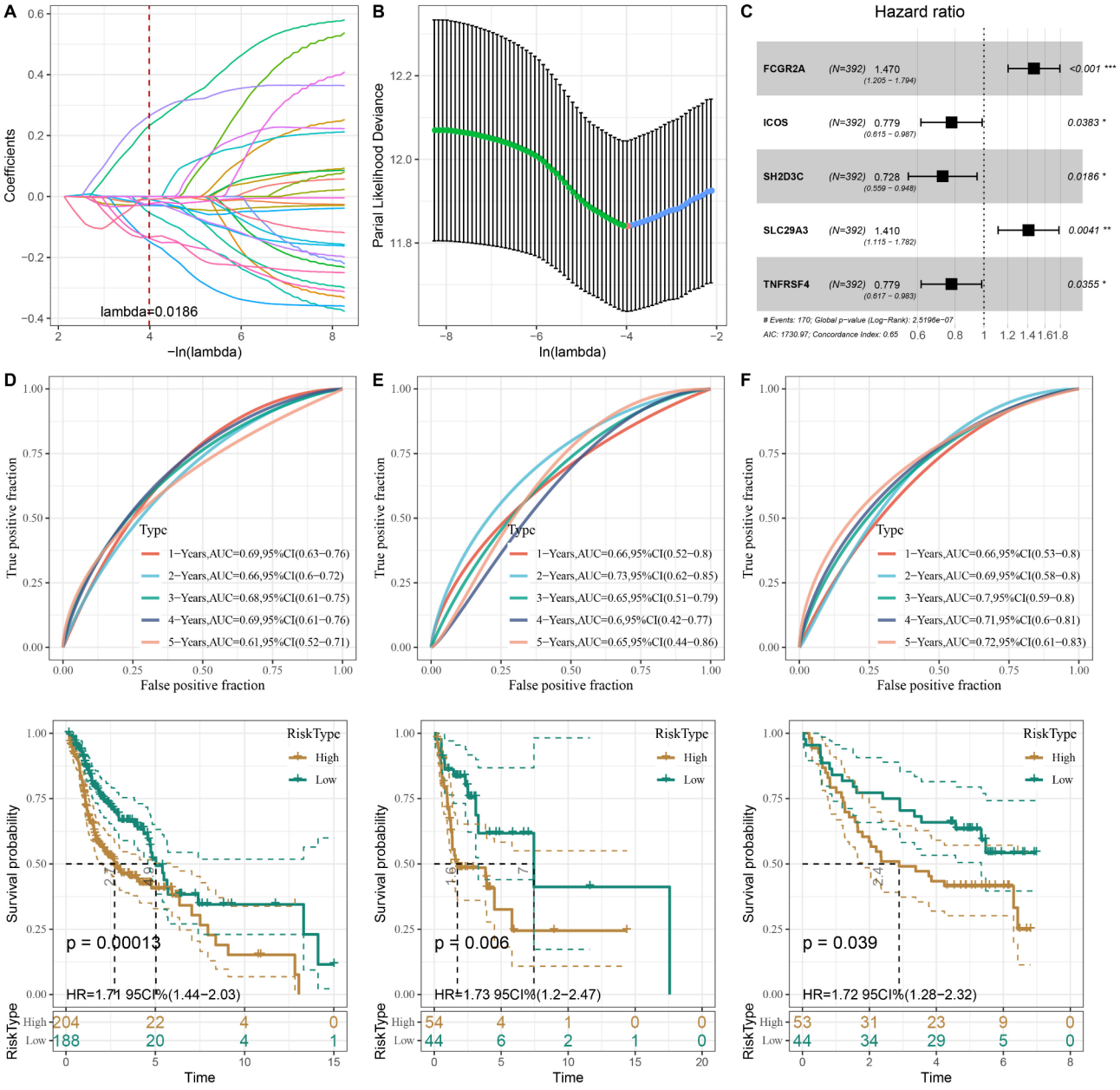

Fig. 5. Establishment and validation of the clinical prognostic model. (**A**) Trajectory of each independent variable as lambda changes. (**B**) Confidence intervals under lambda. (**C**) Forest plot of key genes in the prognostic model. (**D**) ROC curve and KM survival curve of the clinical prognostic model in the TCGA-HNSC training set. (**E**) ROC curve and KM survival curve of the clinical prognostic model in the TCGA-HNSC test set. (**F**) ROC curve and KM survival curve of the clinical prognostic model in the GSE41613 cohort.

Discussion

HNSC arises from the mucosal epithelium of the oral cavity, pharynx, and larynx, representing the most prevalent malignancies in the head and neck region[31]. These carcinomas are characterized by significant heterogeneity and a complex immune microenvironment[32]. Despite considerable advancements in multimodal therapeutic interventions, treatment failure and recurrence remain common, contributing to the majority of mortality cases[33,34]. Immunosuppressive cells play a critical role in the poor survival outcomes observed in patients with HNSC, suggesting that immunotherapy could improve clinical outcomes by counteracting this immunosuppressive environment[35,36]. However, despite the use of multimodal therapeutic strategies, including immunotherapy, a significant proportion of locally advanced tumors eventually recur or progress locally, and response rates to immunotherapy remain suboptimal[33]. The limited effectiveness of current immunotherapeutic approaches underscores the urgent need for further research into novel immune checkpoints and therapeutic strategies to address the pervasive immunosuppression in HNSC[37].

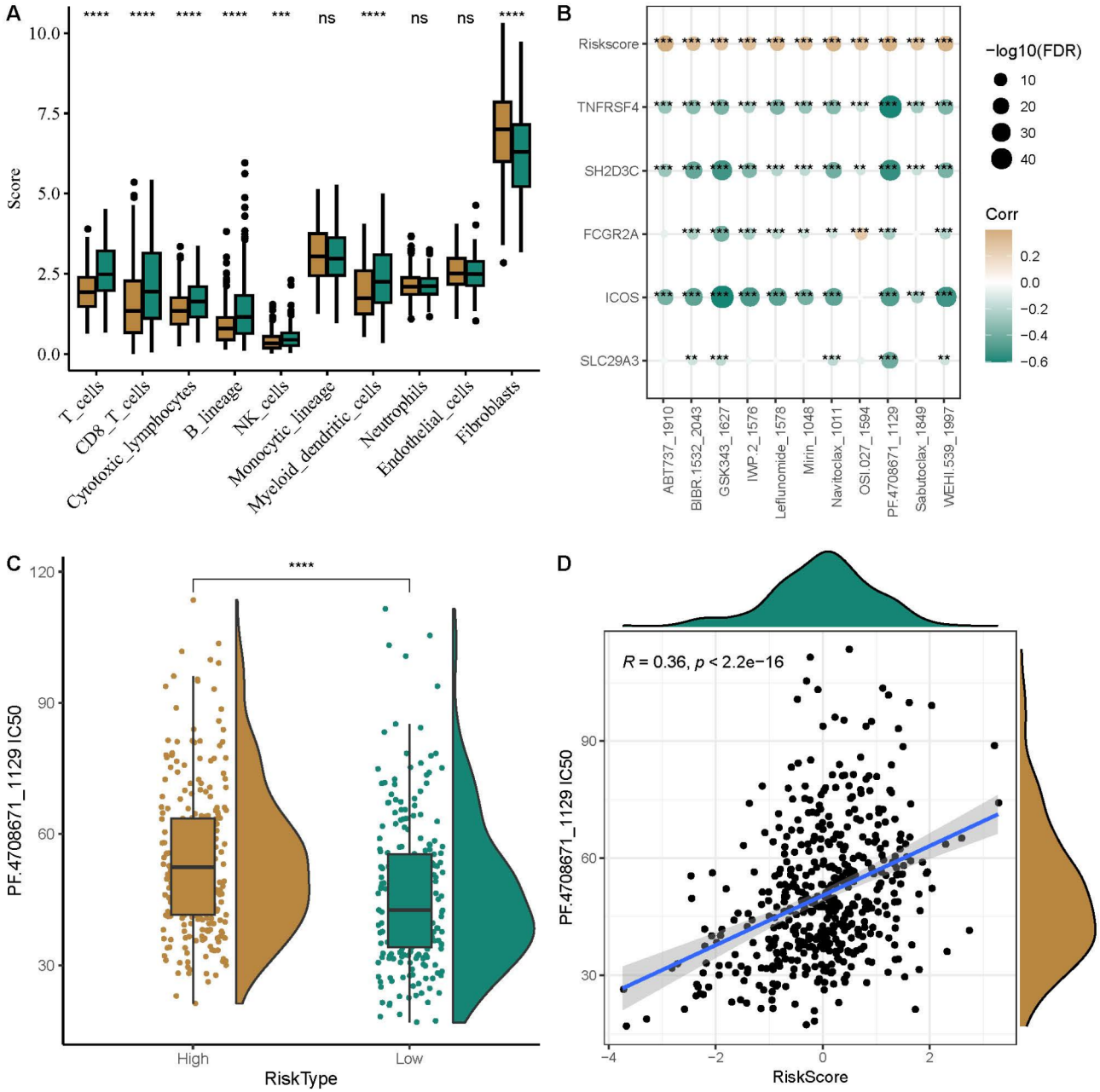

Fig. 6. Risk models in relation to immune microenvironment and chemotherapy drug sensitivity. (**A**) Correlation between RiskScore and immune infiltration levels determined by MCPCounter. (**B**) Correlation analysis between RiskScore, key gene expression in the TCGA-HNSC dataset, and drug IC_{50}. (**C**) Comparison of drug IC_{50} for PF.4708671_1129 between high-risk and low-risk groups. (**D**) Correlation between drug IC_{50} and RiskScore values for PF.4708671_1129.

In this study, clinical data from the UCSC-Xena and GEO databases were used to classify tissue samples into two groups: primary cancer tissue samples and adjacent non-cancerous normal tissue samples. Continued investigation into the role of T cell subpopulations in tumor progression revealed two primary fates for T cells during the transition from normal tissue to HNSC. One subtype, T cell C6, originating from normal tissue, tends to differentiate into a tumor-infiltrating subpopulation. In contrast, T cell C2 emerges at the terminal stage of the differentiation pathway. Throughout this process, immune checkpoints associated with T cell exhaustion, including CTLA4, LAG3, and TIGIT, were progressively upregulated and predominantly enriched within the T cell C2 cluster. TIGIT, an immune checkpoint protein, plays a critical role in normal T cell function[38]. Research indicates that elevated TIGIT expression in CD8 T lymphocytes contributes to CD8 T cell exhaustion in patients with cervical cancer[39]. LAG3, a surface protein interacting with the T Cell receptor complex (TCR, CD3) on both CD4 and CD8 T cells, functions as an inhibitory checkpoint in the immune response[40,41]. These exhausted

T cells lose their ability to effectively recognize and eliminate tumor cells, potentially facilitating tumor immune escape and progression.

Comparative analysis of ligand-receptor pairs between HNSC and normal samples revealed that the LGALS9-CD45 interaction was uniquely present in tumor samples. LGALS9 encodes an S-type lectin involved in key biological processes, including cell adhesion, immune evasion, angiogenesis, and tumor metastasis[42]. CD45, a receptor-associated protein tyrosine phosphatase, is essential for antigen receptor signaling and lymphocyte development[43]. Both LGALS9 and CD45 play significant roles in regulating T cells and promoting immunosuppression[44]. In the development of a clinical prognosis model, five prognostic genes were identified: FCGR2A, inducible co-stimulators (ICOS), SH2D3C, SLC29A3, and TNFRSF4. TNFRSF4, also known as OX40, is a member of the tumor necrosis factor receptor superfamily and acts as a costimulatory molecule expressed by activated immune cells[45]. OX40 is a potent pathway that, when triggered, enhances T cell memory, proliferation, and antitumor activity in patients with metastatic cancer[46]. ICOS, a member of the CD28 family of co-stimulatory molecules, is upregulated on T cells after activation[47]. Research indicates that ICOS signaling influences both effector T cells and regulatory T cells during the anti-tumor response, playing contrasting roles in enhancing immune activation[47,48]. Further investigation into these genes within the prognostic model may clarify their significance in HNSC, offering new molecular targets and therapeutic strategies for clinical management. Additionally, the relationship between the prognostic model and drug sensitivity was explored by determining the maximum half IC_{50} values for each drug within the TCGA-HNSC dataset. This analysis identified 11 drugs significantly associated with RiskScore, which could contribute to the development of personalized cancer treatment approaches.

This study, however, has certain limitations. While the heterogeneity of T cells and the differentiation trajectories of various T cell subsets were explored, the relationships between these distinct subpopulations remain unclear. Further research is needed to elucidate the specific biological processes and molecular mechanisms involved. The findings of this study will serve as a valuable reference for future research aimed at understanding the molecular mechanisms and immune landscape associated with the progression from normal tissue to malignant squamous cell carcinoma.

Conclusion

In conclusion, this study successfully established a single-cell atlas of HNSC and provided a comprehensive analysis of T cell heterogeneity. The differentiation trajectory from normal tissue to HNSC was mapped, revealing that the C2 subgroup of T cells displayed a tendency toward an exhausted state. By intersecting these findings with genes that were notably upregulated in T cell C2, five T cell C2-related signatures were identified through regression analysis, leading to the development of a clinical prognostic model. Additionally, the maximum half IC_{50} values for various drugs were calculated, identifying eleven drugs associated with the risk model, thereby enhancing the potential for personalized cancer therapy. A deep understanding of the primary mechanisms driving T cell exhaustion within the tumor microenvironment is expected to improve the efficacy of immunotherapeutic strategies for HNSC, with significant implications for patient clinical outcomes.

Materials and methods

Data collection and processing

The Fragments Per Kilobase of transcript per Million mapped reads (FPKM) data for the HNSC dataset were sourced from the UCSC Xena database (https://xena.ucsc.edu/). After a rigorous screening process, 490 HNSC samples and 44 normal samples were selected for analysis. Additionally, RNA sequencing (RNA-Seq) data and clinical information from the GSE41613 dataset were obtained from the Gene Expression Omnibus (GEO) database, encompassing 97 HNSC samples. Moreover, scRNA-seq data were retrieved from the GSE181919 dataset, which included 19 tissue samples from 13 patients with HPV-negative HNSC. This collection comprised 13 primary cancer tissue samples (CA) and 6 non-tumor peripheral normal tissue samples (NL).

Data preprocessing

The RNA-seq data from TCGA underwent the following preprocessing steps: Samples without clinical follow-up information were excluded, while those with a survival duration exceeding 30 days were retained. Ensembl identifiers were then converted to corresponding gene symbols, and for genes with multiple symbols, the highest expression level was selected. The expression matrix was reformatted to Transcripts Per Million (TPM) and subsequently transformed using the log2 function.

To preprocess the GEO GSE41613 data, the corresponding chip platform's annotation information was downloaded, and probes were mapped to genes based on this annotation. For genes matched by multiple probes, the average expression value was calculated and used as the gene's expression level.

The preprocessing of the GEO GSE181919 data followed these steps: First, a threshold was set such that each gene had to be expressed in a minimum of three cells, and each cell was required to express at least 200 genes. The single-cell data were filtered, retaining cells with a mitochondrial gene ratio of less than 10% and a gene count between 200 and 6000. The SCTransform function was employed for normalization, followed by principal component analysis (PCA) using the RunPCA function. The harmony package was used to correct for batch effects across different samples[49]. For dimensionality reduction, the first 20 principal components were selected for UMAP analysis. The FindNeighbors and FindClusters functions were then utilized to identify clusters of cell subpopulations. Finally, cell types were annotated using marker genes referenced in PMID:36,828,832 and the Cellmarker 2.0 database.

Construction of single cell pseudotime trajectories

Monocle2 was utilized for the analysis of T cell count data, incorporating their phenotypic information[50]. Cell datasets were constructed using the newCellDataSet function, followed by the exclusion of genes expressed in fewer than 10 cells. Differentially expressed genes between the HNSC group and the normal control group were identified using the differentialGeneTest function. Dimensionality reduction was then performed *via* the reduceDimension function, with parameters set to max_components = 2 and method = 'DDRTree'. The orderCells function was employed to organize the cells and complete the trajectory construction, with additional branches within the normal control group designated as the starting points for the trajectory.

Single-sample gene set enrichment analysis (ssGSEA)

The ssGSEA method, an extension of Gene Set Enrichment Analysis (GSEA), was used to calculate enrichment scores that quantify the absolute gene set enrichment levels for each sample in the dataset[50]. This approach involves comparing individual samples' gene expression profiles against a designated gene set to evaluate the relative enrichment within each sample. The resulting ssGSEA enrichment scores reflect the extent of upregulation or downregulation of specific gene sets in a given sample.

Weighted correlation network analysis (WGCNA)

WGCNA, a systems biology method, was employed to elucidate gene association patterns across samples[51]. This approach enables the identification of gene sets exhibiting strong covariation and facilitates the discovery of potential biomarker genes or therapeutic targets based on the gene sets' intrinsic properties and their relationships with specific phenotypes. In this study, WGCNA was applied to the TCGA HNSC dataset.

Risk model

The identification of genes highly correlated within the T cell C2 module, as determined by WGCNA, resulted in 248 candidate genes. A subsequent univariate Cox regression analysis ($p < 0.05$) narrowed this set to 28 genes with significant prognostic value. To further refine the selection, lasso and stepwise regression techniques were applied, ultimately identifying five genes for the development of a prognostic mode[52]. The resulting risk model was formulated using the equation: $\text{RiskScore} = \Sigma (\beta i \times \text{Expi})$, where "Exp" represents the expression level of the corresponding gene and "β" denotes the associated Cox regression coefficient.

Correlation analysis between risk score and drug sensitivity

The oncoPredict R package was used to estimate drug IC_{50} values in samples from the TCGA-HNSC dataset[53]. Additionally, Spearman correlation analysis was performed to evaluate the relationship between drug sensitivity and RiskScore, with statistical significance defined as $p < 0.05$ and $|\text{cor}| > 0.3$.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon request.

Received: 4 September 2024; Accepted: 19 February 2025
Published online: 10 March 2025

References

1. Leemans, C. R., Snijders, P. J. F. & Brakenhoff, R. H. The molecular landscape of head and neck cancer. *Nat. Rev. Cancer* **18**, 269–282. https://doi.org/10.1038/nrc.2018.11 (2018).
2. Leemans, C. R., Snijders, P. J. F. & Brakenhoff, R. H. Publisher correction: the molecular landscape of head and neck cancer. *Nat. Rev. Cancer* **18**, 662. https://doi.org/10.1038/s41568-018-0057-9 (2018).
3. Baan, R. et al. Carcinogenicity of alcoholic beverages. *Lancet Oncol.* **8**, 292–293. https://doi.org/10.1016/s1470-2045(07)70099-2 (2007).
4. Chaudhary, S. et al. Immunometabolic alterations by HPV infection: new dimensions to head and neck Cancer disparity. *J. Natl Cancer Inst.* **111**, 233–244. https://doi.org/10.1093/jnci/djy207 (2019).
5. Sun, Y. Boron neutron capture therapy: moving towards targeted therapy for locally recurrent head and neck squamous cell carcinoma. *Military Med. Res.* **6**, 32. https://doi.org/10.1186/s40779-019-0224-7 (2019).
6. Chow, L. Q. M. Head and Neck Cancer. *N. Engl. J. Med.* **382**, 60–72. https://doi.org/10.1056/NEJMra1715715 (2020).
7. Zhang, Y. & Zhang, Z. The history and advances in cancer immunotherapy: Understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell Mol. Immunol.* **17**, 807–821. https://doi.org/10.1038/s41423-020-0488-6 (2020).
8. Morad, G., Helmink, B. A., Sharma, P. & Wargo, J. A. Hallmarks of response, resistance, and toxicity to immune checkpoint Blockade. *Cell* **184**, 5309–5337. https://doi.org/10.1016/j.cell.2021.09.020 (2021).
9. Topalian, S. L. et al. Neoadjuvant immune checkpoint Blockade: A window of opportunity to advance cancer immunotherapy. *Cancer Cell.* **41**, 1551–1566. https://doi.org/10.1016/j.ccell.2023.07.011 (2023).
10. Bhatia, A. & Burtness, B. Treating head and neck Cancer in the age of immunotherapy: A 2023 update. *Drugs* **83**, 217–248. https://doi.org/10.1007/s40265-023-01835-2 (2023).
11. Shi, Y. et al. TP53 gain-of-function mutation modulates the immunosuppressive microenvironment in non-HPV-associated oral squamous cell carcinoma. *J. Immunother. Cancer* **11**. https://doi.org/10.1136/jitc-2023-006666 (2023).
12. Elmusrati, A., Wang, J. & Wang, C. Y. Tumor microenvironment and immune evasion in head and neck squamous cell carcinoma. *Int. J. Oral Sci.* **13**. https://doi.org/10.1038/s41368-021-00131-7 (2021).
13. Li, C. et al. Spatial and Single-Cell transcriptomics reveal a Cancer-Associated fibroblast subset in HNSCC that restricts infiltration and antitumor activity of CD8 + T cells. *Cancer Res.* **84**, 258–275. https://doi.org/10.1158/0008-5472.Can-23-1448 (2024).
14. Cheng, D. et al. Molecular and transcriptional basis of bidirectional CD4(+) T cell exhaustion in oropharyngeal squamous cell carcinoma. *MedComm* **5**, e572. https://doi.org/10.1002/mco2.572 (2024).
15. Quah, H. S. et al. Single cell analysis in head and neck cancer reveals potential immune evasion mechanisms during early metastasis. *Nat. Commun.* **14**, 1680. https://doi.org/10.1038/s41467-023-37379-y (2023).

16. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* **15**, 486–499. https://doi.org/10.1038/nri3862 (2015).
17. Baessler, A. & Vignali, D. A. A. T cell exhaustion. *Annu. Rev. Immunol.* **42**, 179–206. https://doi.org/10.1146/annurev-immunol-090222-110914 (2024).
18. Chow, A., Perica, K., Klebanoff, C. A. & Wolchok, J. D. Clinical implications of T cell exhaustion for cancer immunotherapy. *Nat. Rev. Clin. Oncol.* **19**, 775–790. https://doi.org/10.1038/s41571-022-00689-z (2022).
19. Wu, L. et al. Blockade of TIGIT/CD155 signaling reverses T-cell exhaustion and enhances antitumor capability in head and neck squamous cell carcinoma. *Cancer Immunol. Res.* **7**, 1700–1713. https://doi.org/10.1158/2326-6066.Cir-18-0725 (2019).
20. Chapman, N. M., Boothby, M. R. & Chi, H. Metabolic coordination of T cell quiescence and activation. *Nat. Rev. Immunol.* **20**, 55–70. https://doi.org/10.1038/s41577-019-0203-y (2020).
21. Sun, L., Su, Y., Jiao, A., Wang, X. & Zhang, B. T cells in health and disease. *Signal. Transduct. Target. Therapy*. **8**, 235. https://doi.org/10.1038/s41392-023-01471-y (2023).
22. Kersten, K. et al. Spatiotemporal co-dependency between macrophages and exhausted CD8(+) T cells in cancer. *Cancer Cell.* **40**, 624–638e629. https://doi.org/10.1016/j.ccell.2022.05.004 (2022).
23. Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. *Nat. Protoc.* https://doi.org/10.1038/s41596-024-01045-4 (2024).
24. Lv, Y., Ma, X., Ma, Y., Du, Y. & Feng, J. A new emerging target in cancer immunotherapy: Galectin-9 (LGALS9). *Genes Dis.* **10**, 2366–2382. https://doi.org/10.1016/j.gendis.2022.05.020 (2023).
25. Trowbridge, I. S. & Thomas, M. L. CD45: an emerging role as a protein tyrosine phosphatase required for lymphocyte activation and development. *Annu. Rev. Immunol.* **12**, 85–116. https://doi.org/10.1146/annurev.iy.12.040194.000505 (1994).
26. Takefuji, Y. Mitigating biases in feature selection and importance assessments in predictive models using LASSO regression. *Oral Oncol.* **159**, 107090. https://doi.org/10.1016/j.oraloncology.2024.107090 (2024).
27. Liu, Y. et al. Aberrant fragmentomic features of Circulating cell-free mitochondrial DNA enable early detection and prognosis prediction of hepatocellular carcinoma. *Clin. Mol. Hepatol.* https://doi.org/10.3350/cmh.2024.0527 (2024).
28. Xiao, S. et al. Identification of a novel Epithelial-to-mesenchymal-related gene signature in predicting survival of patients with hepatocellular carcinoma. *Comb. Chem. High Throughput Screen.* **25**, 1254–1270. https://doi.org/10.2174/1386207324666210303093629 (2022).
29. Song, X., Dilly, A. K., Kim, S. Y., Choudry, H. A. & Lee, Y. J. Rapamycin-enhanced mitomycin C-induced apoptotic death is mediated through the S6K1-Bad-Bak pathway in peritoneal carcinomatosis. *Cell Death Dis.* **5**, e1281. https://doi.org/10.1038/cddis.2014.242 (2014).
30. Grasso, S. et al. Resistance to selumetinib (AZD6244) in colorectal cancer cell lines is mediated by p70S6K and RPS6 activation. *Neoplasia (New York N Y)* **16**, 845–860. https://doi.org/10.1016/j.neo.2014.08.011 (2014).
31. Johnson, D. E. et al. Head and neck squamous cell carcinoma. *Nat. Reviews Disease Primers* **6**. https://doi.org/10.1038/s41572-020-00224-3 (2020).
32. Mei, Z., Huang, J., Qiao, B. & Lam, A. K. Immune checkpoint pathways in immunotherapy for head and neck squamous cell carcinoma. *Int. J. Oral Sci.* **12**, 16. https://doi.org/10.1038/s41368-020-0084-8 (2020).
33. Bhat, A. A. et al. Tumor microenvironment: an evil nexus promoting aggressive head and neck squamous cell carcinoma and avenue for targeted therapy. *Signal. Transduct. Target. Therapy* **6** https://doi.org/10.1038/s41392-020-00419-w (2021).
34. Szeto, G. L. & Finley, S. D. Integrative approaches to Cancer immunotherapy. *Trends cancer* **5**, 400–410. https://doi.org/10.1016/j.trecan.2019.05.010 (2019).
35. Cohen, E. E. W. et al. The society for immunotherapy of Cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC). *J. Immunother. Cancer*. **7**, 184. https://doi.org/10.1186/s40425-019-0662-5 (2019).
36. Lu, T. et al. CD73 in small extracellular vesicles derived from HNSCC defines tumour-associated immunosuppression mediated by macrophages in the microenvironment. *J. Extracell. Vesicles* **11**, e12218. https://doi.org/10.1002/jev2.12218 (2022).
37. Ruffin, A. T. et al. Improving head and neck cancer therapies by Immunomodulation of the tumour microenvironment. *Nat. Rev. Cancer* **23**, 173–188. https://doi.org/10.1038/s41568-022-00531-9 (2023).
38. Yang, Z. Z. et al. TIGIT expression is associated with T-cell suppression and exhaustion and predicts clinical outcome and Anti-PD-1 response in follicular lymphoma. *Clin. cancer Research: Official J. Am. Association Cancer Res.* **26**, 5217–5231. https://doi.org/10.1158/1078-0432.Ccr-20-0558 (2020).
39. Liu, L. et al. Blocking TIGIT/CD155 signalling reverses CD8(+) T cell exhaustion and enhances the antitumor activity in cervical cancer. *J. Translational Med.* **20**, 280. https://doi.org/10.1186/s12967-022-03480-x (2022).
40. Andrews, L. P., Marciscano, A. E., Drake, C. G. & Vignali, D. A. LAG3 (CD223) as a cancer immunotherapy target. *Immunol. Rev.* **276**, 80–96. https://doi.org/10.1111/imr.12519 (2017).
41. Ruffo, E., Wu, R. C., Bruno, T. C., Workman, C. J. & Vignali, D. A. A. Lymphocyte-activation gene 3 (LAG3): the next immune checkpoint receptor. *Semin. Immunol.* **42**, 101305. https://doi.org/10.1016/j.smim.2019.101305 (2019).
42. Irie, A. et al. Galectin-9 as a prognostic factor with antimetastatic potential in breast cancer. *Clin. cancer Research: Official J. Am. Association Cancer Res.* **11**, 2962–2968. https://doi.org/10.1158/1078-0432.Ccr-04-0861 (2005).
43. Hermiston, M. L., Xu, Z. & Weiss, A. CD45: a critical regulator of signaling thresholds in immune cells. *Annu. Rev. Immunol.* **21**, 107–137. https://doi.org/10.1146/annurev.immunol.21.120601.140946 (2003).
44. Wang, M. et al. Exosomal LGALS9 in the cerebrospinal fluid of glioblastoma patients suppressed dendritic cell antigen presentation and cytotoxic T-cell immunity. *Cell Death Dis.* **11**, 896. https://doi.org/10.1038/s41419-020-03042-3 (2020).
45. Webb, G. J., Hirschfield, G. M. & Lane, P. J. OX40, OX40L and autoimmunity: a comprehensive review. *Clin. Rev. Allergy Immunol.* **50**, 312–332. https://doi.org/10.1007/s12016-015-8498-3 (2016).
46. Porciuncula, A. et al. Spatial mapping and Immunomodulatory role of the OX40/OX40L pathway in human Non-Small cell lung Cancer. *Clin. cancer Research: Official J. Am. Association Cancer Res.* **27**, 6174–6183. https://doi.org/10.1158/1078-0432.Ccr-21-0987 (2021).
47. Simpson, T. R., Quezada, S. A. & Allison, J. P. Regulation of CD4 T cell activation and effector function by inducible costimulator (ICOS). *Curr. Opin. Immunol.* **22**, 326–332. https://doi.org/10.1016/j.coi.2010.01.001 (2010).
48. Metzger, T. C. et al. ICOS promotes the function of CD4 + Effector T cells during Anti-OX40-Mediated tumor rejection. *Cancer Res.* **76**, 3684–3689. https://doi.org/10.1158/0008-5472.Can-15-3412 (2016).
49. Giuliani, A. The application of principal component analysis to drug discovery and biomedical data. *Drug Discovery Today* **22**, 1069–1076. https://doi.org/10.1016/j.drudis.2017.01.005 (2017).
50. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982. https://doi.org/10.1038/nmeth.4402 (2017).
51. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**. https://doi.org/10.1186/1471-2105-9-559 (2008).
52. Koletsi, D. & Pandis, N. Survival analysis, part 3: Cox regression. *Am. J. Orthod. Dentofac. Orthopedics: Official Publication Am. Association Orthodontists its Constituent Soc. Am. Board. Orthod.* **152**, 722–723. https://doi.org/10.1016/j.ajodo.2017.07.009 (2017).
53. Yang, W. et al. Genomics of drug sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **41**, D955–961. https://doi.org/10.1093/nar/gks1111 (2013).

Acknowledgements
These results in whole or part are based on data generated by TCGA Research https://www.cancer.gov/tcga.

Author contributions
Jie Liu and Yuanyuan Zhang wrote the manuscript. Penghui Li and Lian Zheng designed the manuscript.

Funding
This study was supported by the Henan Provincial Medical Science and Technology Research Plan (SBGJ202102175).

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-91299-z.

Correspondence and requests for materials should be addressed to L.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.