RESEARCH



High-dimensional deconstruction of HNSC reveals clinically distinct cellular states and ecosystems that are associated with prognosis and therapy response



Lei Xiao^{1,2,3,4}, Zhe Shen^{1,2,3,4}, Zhaoyu Pan^{1,2,3,4}, Yuanzheng Qiu^{1,2,3,4}, Donghai Huang^{1,2,3,4}, Yong Liu^{1,2,3,4}, Chao Liu^{1,2,3,4*} and Xin Zhang^{1,2,3,4*}

Abstract

Background Characterizing the variety of cell types in the tumor microenvironment (TME) and their organization into cellular communities is vital for elucidating the biological diversity of cancer and informing therapeutic strategies.

Methods Here, we employed a machine learning-based algorithm framework, EcoTyper, to analyze single-cell transcriptomes from 139 patients with head and neck squamous cell carcinoma (HNSC) and gene expression profiles from 983 additional HNSC patients, aiming to delineate the fundamental cell states and ecosystems integral to HNSC.

Results A diverse landscape of 66 cell states and 9 ecosystems within the HNSC microenvironment was identified, revealing classical cell types while also expanding upon previous immune classifications. Survival analysis revealed that specific cell states and ecotypes (ecosystems) are associated with patient prognosis, underscoring their potential as indicators of clinical outcomes. Moreover, distinct cell states and ecotypes exhibited varying responses to immuno-therapy and chemotherapy, with several showing promise as predictive biomarkers for treatment efficacy.

Conclusion Our large-scale integrative transcriptome analysis provides high-resolution insights into the cellular states and ecosystems of HNSC, facilitating the discovery of novel biomarkers and supporting the development of precision therapies.

*Correspondence: Chao Liu

- liuchaoent@csu.edu.cn Xin Zhang
- xin Znang xinzhang@csu.edu.cn
- ¹ Department of Otolaryngology Head and Neck Surgery, Xiangya
- Hospital, Central South University, 87 Xiangya Road, Changsha 410008, Hunan, China ² Otolaryngology Major Disease Research Key Laboratory of Hunan
- Province, Changsha 410008, Hunan, China
- ³ Clinical Research Center for Pharyngolaryngeal Diseases and Voice Disorders in Hunan Province, Changsha 410008, Hunan, China

⁴ National Clinical Research Center for Geriatric Disorders (Xiangya

Hospital), Changsha 410008, Hunan, China

Introduction

Head and neck squamous cell carcinoma (HNSC) is one of the most common malignant tumors worldwide, with approximately 800,000 new cases reported each year. Its prognosis remains poor,with a 5 year survival rate of less than 50% [1]. Despite recent advances in treatment strategies, particularly targeted therapies [2] and immune checkpoint therapies [3], significant improvements in patient prognosis have remained elusive [4]. The poor prognosis of HNSC patients can primarily be attributed to the marked heterogeneity of the disease, which often results in suboptimal treatment [4–6]. Therefore, early characterization of the tumor's



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

heterogeneity is essential for guiding precise treatment strategies and improving patient prognosis.

The American Joint Committee on Cancer (AJCC) classification is a key clinical tool for assessing risk and guiding treatment decisions for HNSC patients [7]. However, the current AJCC classification has notable limitations, as it considers only clinical stage and overlooks the tumor heterogeneity among patients at the same stage. Consequently, this can lead to inadequate clinical decision-making, potentially resulting in either overtreatment or undertreatment. Therefore, a more comprehensive understanding of the tumor's immune microenvironment and its heterogeneity is crucial for informing targeted therapies and improving clinical outcomes in HNSC patients [4].

The rapid advancement of single-cell sequencing technology has revolutionized our ability to profile individual cells with unprecedented resolution, providing deeper insights into the tumor microenvironment [8– 10]. Ruben Bill et al. demonstrated that SPP1/CXCL9 serves as a macrophage marker and capable of predicting HNSC prognosis based on single-cell sequencing data from 52 samples [11]. Similarly, Z. L. Liu et al. reported that interactions between SPP1 + macrophage and POSTN+fibroblast promote tumor progression and are associated with poor prognosis in HNSC patients¹². Previous studies have primarily focused on the relationship between specific cell types and patient outcomes, which lack a comprehensive understanding of the diverse cellular states within the tumor microenvironment of HNSC and their combined impact on patient prognosis. This gap underscores the need for an integrated perspective that accounts for the complex interplay of cellular components within the tumor context.

In this study, we employed the EcoTyper machine learning framework [13] to identify the fundamental cell states and communities within HNSC, revealing 66 cell states and 9 ecosystems that expand upon previous immune classifications. We validated these findings using multicolor immunofluorescence, spatial transcriptomics, single-cell, and bulk RNA transcriptomics, and extended our analysis to a broader pan-cancer context. Survival analysis highlighted the prognostic significance of distinct cell states and ecosystems, demonstrating significant correlations with patient outcomes. Moreover, we investigated the potential application of these cell states and ecosystems in treatment selection, finding that different cell states and ecosystems exhibited distinct responses to immunotherapy and chemotherapy, with multiple cell states and multicellular communities showing potential as predictors for treatment response.

Results

Identification and validation of HNSC cell states

EcoTyper employs a machine learning framework to extract cell type-specific gene expression profiles from the transcriptome, identify transcriptional cell states within each cell type, and define a tumor ecosystem ("ecotype") composed of coexisting cell states [13] (Fig. 1). To identify and validate HNSC specific cell states, as well as multicellular communities, and to explore their clinical significance, we gathered data from multiple sources, including 131 single-cell sequencing samples of HNSC from 7 GEO cohorts (104 primary samples in discovery cohort and 27 metastatic samples in validation cohort), 7 BLCA single-cell sequencing samples from GSE135337, 11 melanoma (SKCM) single-cell samples from GSE215121, and 8 HNSC spatial transcriptomic samples from GSE181300. In addition, we obtained microarray data from 441 HNSC patients across three GEO cohorts, RNA-seq data from 519 HNSC patients in the TCGA-HNSC cohort, 472 patients across three publicly available immunotherapy cohorts, as well as RNAseq data from 23 patients in our Xiangya hypopharyngeal carcinoma chemotherapy cohort. (Fig. 1, Figure S1, Supplementary Table 1).

We initially performed quality control and clustering annotation on 104 primary HNSC single-cell samples, identifying 256,898 high-quality cells classified into 13 distinct cell types (Fig. 2A). Each cell type was annotated based on classic HNSC marker genes [11, 14–17] (Fig. 2B, Figure S2A). We then applied the EcoTyper framework with non-negative matrix factorization (NMF) to identify between 3 and 9 cell states for each cell type, ultimately identifying 66 different cell states (Fig. 2C, Supplementary Table 2). These include: (1) B Cell: Six states, including germinal center B cell expressing LRMP and AICDA, and naïve B cells expressing IGHD and FCER2 [18]. (2) Plasma Cell: Nine states, including ATP5MC1+plasma cell and ATPIF1+plasma cell. (3) CD4+T cell: Three states, including naïve CD4 T cell states expressing CCR7. (4) CD8+T cell: Three states, including GZMA+effector CD8 T cell and FOXP3+regulatory CD8 T cell. (5) NK Cell: Five states, including NKT cells with high expression of IFNG. (6) Monocyte: Three states, including intermediate monocytes expressing both CD14 and FCGR3A (7) Macrophage: Seven states, including SPP1+macrophage, CXCL9+macrophage, APOE + macrophage, TGFBI + CXCL8 + macrophage [19], and proliferative macrophage. (8) cDC: Three major cDC states, including cDC1 expressing CLEC9A and cDC2 expressing CLEC10A. (9) pDC: Six major pDC states. (10) Mast Cells: Five major states. (11) Fibroblast: Four states, including



Fig. 1 Machine learning framework for large-scale identification and validation of HNSC cell states and ecosystems This schematic diagram illustrates the application of EcoTyper in HNSC patients. Initially, cell states were identified in a discovery cohort composed of single-cell sequencing data from multiple cohorts of primary HNSC patients. These cell states were subsequently validated using metastatic single-cell sequencing samples, and associations between cell state abundance and patient outcomes were analyzed. HNSC ecotypes were identified by examining co-occurrence patterns among cell states, and spatial transcriptomics was employed to validate the spatial distribution of both ecotypes and cell states. Finally, the association between tumor ecotypes and patient outcomes was examined, along with the analysis of the relationship between cell states, ecotypes, and treatment sensitivity across several immunotherapy and chemotherapy cohorts

TGFBI+WNT5A+fibroblast, ACTA2+fibroblast, and CFD+inflammatory fibroblast. (12) Endothelial Cell: Six distinct states, including lymphatic endothelial cell (characterized by CCL21 and PROX1 expression), arterial endothelial cell (characterized by HEY1 and IGFBP3 expression), and venous endothelial cell (characterized by ACKR1 expression). (13) Epithelial Cell: Six states, including a highly malignant epithelial subtype characterized by high expression of WNT5A.

To validate the cell states identified from primary HNSC single-cell samples, we used the same clustering and annotation strategy to classify 27 metastatic HNSC samples into the same 13 cell types (Figure S2B, C). Using the EcoTyper framework, we successfully restored nearly all cell states identified in the primary samples (Fig. 2D, Supplementary Table 3), demonstrating their robustness.

To further validate the accuracy and reliability of the classifications, we performed additional analyses with the ROGUE and scSHC R packages. The ROGUE index was calculated for each sample and aggregated by cell type, confirming high purity of the identified cell states (Figure S2D). Furthermore, the scSHC testClusters function showed high consistency with the original EcoTyperbased classification, further validating the clustering results (Supplementary Table 4). These complementary

analyses reinforce the reliability and robustness of the identified cell states.

Deciphering the prognostic landscape of HNSC cell states

Cell states have been shown to have the ability to predict patient survival [13, 20, 21]. However, their prognostic significance in HNSC remains unclear. To address this gap, we leveraged the unique capabilities of EcoTyper to map the prognostic profiles of 66 cell states in HNSC. Analysis of overall survival dichotomized these cell states into favorable and unfavorable categories, identifying six cell states linked to poor survival and seven significantly associated with favorable outcomes (Fig. 3A). Notably, the association between cell state abundance and patient outcomes was consistent in the validation cohort (Fig. 3B). Among the 13 cell states significantly associated with patient outcomes in the training cohort, six maintained their significance in the validation cohort. Additionally, the direction of association was preserved for most cell states, demonstrating strong correlations across cohorts despite differing platforms used for bulk gene expression analysis (Figure S3A).

In line with previous findings [11, 22], a higher abundance of SPP1+M2-like immunosuppressive macrophage (S06) was associated with poorer outcomes



Fig. 2 Discovery and validation of HNSC cell states **A** UMAP plot of the discovery cohort composed of primary HNSC samples. **B** Expression of marker genes for 13 cell subpopulations in the discovery cohort. **C** Heatmap showing distinct cell states across 13 cell types. **D** Recovery of cell states in metastatic single-cell HNSC samples, with a z-score > 1.65 considered significant

in both the training and validation cohorts (Fig. 3C). We also identified several additional prognostic associations. For instance, ATPIF1+ATP5G1+mitochondrial metabolism-related plasma cell (S08) were linked

to improved outcomes in HNSC patients across both cohorts (Fig. 3C). Furthermore, we compared the prognostic significance of clinical stage with that of individual cell states in HNSC patients. Clinical stage exhibited a more stable prognostic discriminatory effect compared to individual cell states; however, certain cell states, such as ACTA2+fibroblast (S02), demonstrated robust prognostic efficacy across all cohorts (Figure S3B). To further validate these findings and provide a comprehensive visualization of ACTA2 + fibroblast, we analyzed our Xiangya HNSC TMA cohort to identify ACTA2+fibroblast using two markers, ACTA2 and COL1A1 (Fig. 3D, Figure S3C, D). In Fig. 3E, we further explored the relationship between ACTA2+fibroblast abundance and prognosis in both the publicly available datasets and the Xiangya HNSC TMA cohort. Consistently, a higher abundance of ACTA2 + fibroblast was correlated with poorer prognosis across all cohorts analyzed. Additionally, we explored the relationship between cell states associated with better or worse prognosis and HPV infection using data from the TCGA database. We found that cell states linked to a relatively favorable prognosis were more abundant in HPVpositive patients, whereas those associated with a poorer prognosis were more prevalent in HPV-negative patients (Figure S3E, F).

Reconstructing HNSC cellular communities

Next, we employed the EcoTyper framework to analyze patterns of cell state co-occurrence and reconstruct the core cellular communities within HNSC. This analysis revealed nine distinct "HNSC ecotypes", each comprising 3 to 12 co-occurring cell states (Fig. 4A, B). Notably, almost every HNSC patient exhibited a dominant ecotype, with many tumors comprising multiple ecotypes (Fig. 4A). We also identified these ecotypes in metastatic HNSC single-cell samples, as well as in TCGA and GEO transcriptomic datasets, observing consistent restoration of the majority of ecotypes across different sample types (Fig. 4C). Besides, E5 and E9 were scarcely detected in metastatic HNSC single-cell samples, likely due to the limited sample size of only 27 metastatic HNSC single-cell samples (Fig. 4D).

To further investigate the colocalization and interactions between cell states within each ecotype and to validate the nine ecotypes identified by EcoTyper, we analyzed eight HNSC spatial transcriptomic samples. We performed reference-guided recovery to estimate the abundance of each cell state and ecotype within spatial barcodes. Cell states belonging to the same ecotype exhibited stronger spatial correlations than those from different ecotypes (Fig. 4E) and were frequently located within the same tumor region (Fig. 4F, Figure S4A, B). Interestingly, different ecotypes varied in abundance across distinct tumor regions (Figure S4C). In the nine tumors analyzed by spatial transcriptomics, most ecotypes displayed significant spatial clustering (Fig. 4G), suggesting that these ecotypes represent distinct functional units comprising various cell states in HNSC. Moreover, we investigated the intercellular signaling networks within the HNSC ecotypes using CellChat cell interaction analysis (Figure S4D, Supplementary Table 5). In the E7 ecotype, which is associated with a favorable prognosis, NKT cells (NK_cell_S02) exhibited strong interactions with cDCs and pDCs through CD8 receptor signaling. Pathways such as CXCL9-CXCR3 and MIF-CD74_CXCR4 suggest that these interactions activate immune responses, thereby promoting antitumor immunity. In contrast, the E8 ecotype, associated with a poor prognosis, demonstrated a network dominated by TGF β signaling. In this ecotype, fibroblasts (Fibroblast_S01) interacted with endothelial and epithelial cells via TGFB1-TGFBR1/2 signaling, fostering a pro-tumorigenic and immunosuppressive microenvironment. These results highlight key cellular interactions that may play a pivotal role in shaping the HNSC microenvironment.

Characteristics of the multicellular ecosystem in HNSC

After identifying the nine major multicellular ecosystems in cancer, we explored their clinical characteristics (Fig. 5A). In the TCGA-HNSC cohort, two ecotypes were significantly associated with prognosis. E7, primarily composed of anti-tumor immune cells such as cDC1 and NKT cells, is characterized by the activation of immune-related pathways, including T cell receptor signaling, NK-mediated immunity, and leukocyte activation. This ecotype was strongly associated with favorable prognosis in HNSC patients. In contrast, E8, which is enriched in cell adhesion and growth-related pathways, was closely associated with a higher risk of death and marked by elevated levels

(See figure on next page.)

Fig. 3 Association of cell state abundance with patient outcomes across different groups **A** Survival associations of 66 cell states in the TCGA-HNSC cohort, with the most significant marker genes for both unfavorable and favorable cell states indicated. If a cell state is associated with shorter survival, the survival association is represented as the -log10(p-value) multiplied by 1. For cell states associated with favorable outcomes, it is shown as -1. **B** Scatter plot illustrating the correlation between cell state survival associations in the TCGA-HNSC cohort and the GEO integrated cohorts (GSE41613, GSE42743, GSE65858). Spearman correlation coefficients and two-sided P-values are indicated. The graph also displays the linear regression best-fit line and the 95% confidence interval. **C** Kaplan-Meier survival curves for macrophages (S06) and plasma cells (S08) in the TCGA-HNSC cohort. **D** Representative multicolor immunofluorescence image of ACTA2 + fibroblast. **E** Prognostic curves for ACTA2 + fibroblast





Fig. 3 (See legend on previous page.)



Fig. 4 Discovery of multicellular communities in HNSC A Heatmap of tumor ecotypes based on different combinations of cell states. B Network diagram illustrating ecotype composition, with the width of each edge representing the Jaccard index from the discovery cohort. C Heatmap of tumor ecotypes across various validation cohorts. D Proportion diagram showing the distribution of tumor ecotypes in different cohorts. E Violin plots comparing Spearman spatial correlation between cell states of different ecotypes and within the same ecotype. F Spatial distribution of each cell state within the E8 ecotype, analyzed using spatial transcriptomics. G Spatial clustering of HNSC ecotypes measured using Moran's I in 8 spatial transcriptomic samples of HNSC. A z-score > 1.96 was considered significantly higher than the degree of clustering expected by chance

of TGFBI+fibroblasts and TGFBI+macrophages. To explore the genomic basis of prognostic differences between the E7 and E8 ecotypes, we analyzed the types of genomic alterations in the TCGA-HNSC cohort. The E8 ecotype exhibited more frequent alterations, particularly in genes related to cell adhesion and growth, which could contribute to the poorer prognosis observed in E8 (Fig. 5B, C).

Using the TCGA-HNSC cohort, we further compared the prognostic discrimination efficiency of ecotypes and clinical staging. E7 and E8 demonstrated prognostic discrimination nearly equivalent to that of clinical staging

(Figure S5A). Moreover, the combination of N-stage and ecotype abundance achieved the most effective prognostic stratification, suggesting that tumor ecotypes provide a unique perspective in predicting clinical outcomes (Figure S5A). Additionally, compared to individual cell states, tumor ecotypes exhibited stronger consistency in prognostic stratification across the TCGA and GEO cohorts (Figure S5B). Besides, we observed that the abundance of E7 and E8 ecotypes did not show strong correlations with traditional clinicopathological parameters such as TNM staging or clinical stage. However, interestingly, ecotype abundance did show some correlation with pathological grading, though this was primarily observed in the G4 group, which had a relatively small sample size (Figure S5C-D). Despite this, these findings underline that ecotype classification is an independent prognostic factor, providing valuable information that goes beyond traditional clinical metrics. This highlights the potential of tumor ecotypes as complementary tools in clinical prognostication and their ability to offer additional insights into tumor progression and patient outcomes.

From the Sankey diagram in Fig. 5D, we observed that the tumor ecotypes defined in this study differ substantially in composition from previous pan-cancer ecotype classifications [13], though some similarities are evident. In this study, E7 is associated with a favorable prognosis, with most E7 samples of HNSC aligning with CE9 and CE10 in pan-cancer ecotype, which were also linked to the best prognoses in previous pan-cancer classifications. In contrast, E8 is associated with a poor prognosis, and roughly half of the E8 samples from HNSC correspond to CE1, which had the worst prognosis in the prior classification. Besides, we found that the E7 and E8 ecotypes offer superior prognostic prediction capabilities compared to the previous pan-cancer ecotype classification. As shown in Fig. 5E, E7 and E8 outperform CE1 and CE10 in predicting patient outcomes across multiple cohorts, demonstrating the enhanced accuracy of this classification in guiding prognostic decisions for HNSC. These findings suggest that our tumor ecotype classification, derived from single-cell HNSC samples, sheds new light on the tumor microenvironment in HNSC.. Furthermore, E8 patients are predominantly HPV-negative, which may indicate reduced sensitivity to immunotherapy.

Prediction of immunotherapy response using HNSC cell states and ecotypes

Next, we investigated whether ecotypes could predict immunotherapy response. We first analyzed BLCA and melanoma single-cell cohorts within the ecotype framework. Our analysis revealed that the common cell types in both single-cell sequencing cohorts were largely restored (Figure S5E-F, Supplementary Table 3), suggesting that the cell state and ecotype analysis derived from HNSC samples can be effectively applied to other cancer types. Building on this, we collected tumor expression data from 472 patients with advanced cancers who had received immune checkpoint blockade therapies, anti-PDL1 (urothelial carcinoma), anti-PD1 (melanoma), or anti-CTLA4 (melanoma) for an integrated analysis. To quantify the performance, we assessed the continuous association of ecotypes with overall survival in patients receiving immunotherapy, as well as their binary association with immunotherapy response. E7, which is enriched in NKT cell and cDC1, demonstrated the best response to immunotherapy compared to other ecotypes, while E8 showed relative insensitivity to treatment (Fig. 5F). We also compared the performance of ecotypes with 77 candidate biomarkers, including 66 cell states identified by EcoTyper, 3 published immune checkpoint inhibitor (ICI) response signatures [23-25], and commonly used ICI response marker genes such as CD274, PDCD1, and CTLA4 [26]. Notably, the abundance of E7 outperformed nearly all other indicators, including those specifically designed to predict ICI response (Fig. 5F; Supplementary Table 6), highlighting its strong association with favorable immunotherapy outcomes. Conversely, patients with high levels of SPP1+macrophage (S06) and E8 showed poor responses to immunotherapy.

Besides, we analyzed pathological sections from two immunotherapy-resistant HNSC patients, further supporting the presence of the E8 ecotype. As shown in Fig. 5G, both patients exhibited highpower fields with the co-existence of TGFBI+macrophages and TGFBI+fibroblast in the multicolor

(See figure on next page.)

Fig. 5 HNSC ecotype characteristics and their association with immunotherapy response A Ecotype characteristics in the discovery cohort. Top: Ecotype-specific survival associations TCGA-HNSC cohort with favorable (blue) or unfavorable (red) survival outcomes indicated. Middle: Average abundance of each cell type within each ecotype. Bottom: Gene functional enrichment analysis for the 9 identified ecotypes. B–C Plot of genomic alterations from TCGA in E7 and E8 ecotype. D Sankey plots depicting the associations between ecotypes identified in this study and pan-cancer ecotypes, as well as HPV status. E A comparison of the prognosis discrimination ability between the ecotypes identified in this study and pan-cancer ecotypes. F Associations of 86 signatures with overall survival and response to immune checkpoint inhibitors (ICIs) in 472 patients with advanced melanoma or BLCA. G Representative images showing the coexistence of TGFBI + macrophages and TGFBI + fibroblasts within the E8 ecotype in the multi-color immunohistochemical slices of two immunotherapy-resistant patients, observed at 40 × magnification



Fig. 5 (See legend on previous page.)

immunohistochemistry sections of their HNSC tissues. These findings demonstrate that multicellular ecosystems can capture signals related to immunotherapy response within the tumor immune microenvironment of HNSC, which are highly predictive of immunotherapy outcomes.

Prediction of chemotherapy response using HNSC cell states and ecotypes

TPF treatment regimens play a crucial role in managing HNSC and are often combined with immunotherapy to enhance tumor control. The tumor microenvironment significantly influences the response of solid tumors to systemic therapies, such as immunotherapy and chemotherapy. To further explore its impact on prognosis and chemotherapy sensitivity, we conducted an integrated analysis on 78 patients from the TCGA-HNSC cohort who received TPF monotherapy or combination therapy, along with 23 HNSC patients treated with TPF chemotherapy at Xiangya Hospital. Our findings revealed that patients with a high abundance of the E1 ecotype responded best to chemotherapy (Fig. 6A). In both the Xiangya and TCGA cohorts, the abundance of E1 was significantly elevated in chemotherapy-sensitive patients (Fig. 6B). Compared to the previous pan-cancer ecotype classification, the predictive efficacy of E1 for chemotherapy sensitivity was notably stronger, suggesting the superior ability of this classification in guiding treatment decisions. Conversely, patients with a high abundance of ACTA2+fibroblast (S02) exhibited the worst prognosis in response to chemotherapy (Fig. 6A). In these same cohorts, ACTA2+fibroblast (S02) abundance was markedly reduced in chemotherapy-sensitive patients (Fig. 6C). Moreover, ROC curve analysis indicated that both E1 and fibroblast (S02) exhibited strong predictive efficacy for chemotherapy sensitivity (Figure S6A, B). Survival analysis further indicated that patients with high levels of fibroblast (S02) had significantly worse prognoses in both cohorts (Figure S6C, D). Additionally, we further validated the impact of ACTA2+fibroblast (S02) abundance on chemotherapy sensitivity in the TMA chemotherapy cohort, confirming that high levels of ACTA2+fibroblasts consistently indicated chemotherapy resistance across all cohorts (Fig. 6D, E and Supplementary Table 7). These results show that HNSC cell states and ecotypes may be helpful for predicting the chemotherapy response.

Methods

Human subjects

This study collected tissue samples from 122 patients with HNSC at Xiangya Hospital, Central South University. Among them, 23 fresh tissue samples were obtained from the patients who underwent TPF chemotherapy prior to chemotherapy. Fresh tissue specimens were collected at the time of biopsy, and total RNA was immediately extracted for RNA-seq. Paraffin-embedded samples were collected during biopsy or surgical procedures and immediately preserved in formalin. Among these patients, 26 received induction or concurrent TPF chemotherapy, 2 advanced HNSC patients received immunotherapy (anti-PD-1). Clinical specimens were collected from January 2020 to December 2023, alongside recording patients' clinical information and prognoses. Followups were conducted every 3 to 6 months until death or loss to follow-up. The study was conducted in accordance with the Declaration of Helsinki. The study protocol was approved by the Ethics Review Committee of Xiangya Hospital, Central South University. Written informed consent was obtained from all participants, and all samples and data collected in this study were handled in compliance with ethical guidelines.

TMA construction

Two pathologists visually inspected HE-stained sections to evaluate and select representative formalin-fixed paraffin-embedded (FFPE) tissue blocks from 97 HNSC patients, including 31 pairs of tumor and adjacent normal tissue blocks. Based on HE-guided positioning, 1.5 mm cores were extracted from the tumor areas using a tissue microarray sampling tool. The cores were then inserted into recipient paraffin blocks according to a predetermined array layout. A tissue microarray fusion instrument was used to repeatedly fuse the donor tissue cores and recipient wax block to ensure complete integration, forming the tissue microarray module. Finally, the microarray was sectioned into 3 μ m slices for multi-color immunofluorescence staining.

Multicolor immunofluorescence staining

Paraffin sections were baked at 60 °C for 1-2 h to ensure complete fixation. The sections were then dewaxed and hydrated by sequential immersion in xylene I and II, followed by treatment with graded ethanol to preserve sample integrity. After several washes in sterile water, antigen retrieval was performed using EDTA antigen retrieval buffer (pH 8.0; Powerful Biology, B0035). Sections were heated in a microwave at high temperature for 3.5 min, then at low temperature for 15 min to ensure optimal antigen exposure. Next, sections were incubated with 3% H₂O₂ at room temperature for 25 min to inactivate endogenous peroxidase, followed by incubation with 3% BSA for 10 min to block nonspecific binding sites. Primary antibody incubation was carried out, followed by HRP-conjugated secondary antibodies and TSA (tyramide signal amplification) conjugated fluorescent dye, repeating the staining cycle three times to label all



Fig. 6 Association between HNSC cell states, ecotypes, and chemotherapy response A Associations between the 66 cell states, the ecotypes identified in this study, pan-cancer ecotypes, and chemotherapy response were examined. B Correlation between E1 ecotype abundance and chemotherapy sensitivity in the Xiangya and TCGA chemotherapy cohorts. C Correlation between fibroblast (S02) abundance and chemotherapy sensitivity in the Xiangya and TCGA chemotherapy cohorts. D Representative multicolor immunofluorescence images showcasing varying abundances of ACTA2 + fibroblasts in the TMA chemotherapy cohorts. E Correlation between ACTA2 + fibroblast abundance and chemotherapy sensitivity in TMA chemotherapy cohorts

targets. After each cycle, antigen retrieval was repeated using EDTA buffer and microwave heating (3.5 min at high temperature, 15 min at low temperature). The primary antibodies used to verify ACTA2+fibroblast localization were as follows: anti-human ACTA2 (Abcam, ab124964; 1:1000), and anti-human COL1A1 (Abcam, ab138492; 1:1500). Multispectral imaging of the stained TMAs was conducted using the Vectra 3.0 system (Akoya), and images were analyzed with inForm software (Akoya). In this study, a total of 94 tissue samples with high staining quality and suitable for analysis were obtained (Figure S3D, Supplementary Table 7).

Public dataset cohorts

We collected gene expression data from a total of 960 patients with HNSC from the TCGA and GEO databases.

The three GEO datasets (GSE42743, GSE41613, and GSE65858) were preprocessed using the robust multiarray average (RMA) algorithm implemented in the "affy" package in R to ensure consistent normalization across microarray data. To mitigate potential batch effects that could arise from combining different datasets, we applied the "sva" package in R, specifically using the Combat function to harmonize expression values. Probes mapping to multiple genes were removed to prevent ambiguous gene assignments, and for genes represented by multiple probes, the median expression value was selected to maintain a singular representative measure per gene.

Additionally, we obtained three independent immunotherapy cohorts from the Tumor Immunotherapy Gene Expression Resource (http://tiger.canceromics.org/) database. To ensure consistency in our analyses, only patient samples collected prior to immunotherapy treatment were included, eliminating the potential confounding effects of therapy-induced gene expression changes.

For single-cell RNA sequencing (scRNA-seq) data, we compiled samples from seven GEO datasets, while excluding nasopharyngeal carcinoma and lymph nodederived samples to maintain specificity to HNSC. In total, 131 samples were retained for further analysis. To minimize technical artifacts, we implemented several rigorous quality control steps. The python package "scrublet" was used to detect and remove doublets, thereby reducing artificial cell multiplets that could confound downstream clustering and annotation. Data preprocessing was carried out using Seurat (v5.0), including variable gene selection, integration of multiple samples, dimensionality reduction, clustering, and differential expression analysis. To ensure the reliability of our dataset, we applied strict filtering criteria: (i) Cells with mitochondrial RNA content exceeding 10% were removed to eliminate dying or stressed cells. (ii) Cells with fewer than 1000 or more than 5000 detected genes were excluded to avoid lowquality or doublet cells. (iii) Cells with over 75,000 unique molecular identifiers (UMIs) were removed to prevent potential artifacts from over-amplification.

For the remaining cells, we used the 'NormalizeData' function in Seurat to standardize expression values and 'ScaleData' to mean-center and scale gene expression, ensuring comparability across samples. To correct for batch effects, the HarmonyIntegration method was employed, further improving the consistency of the integrated dataset. Clustering was performed using the FindClusters function with a resolution parameter set to 1.2. Cell annotation was initially conducted using the SingleR algorithm, providing a preliminary classification based on reference transcriptomic profiles, and was subsequently refined manually using classic cell marker genes to enhance accuracy.

Furthermore, we incorporated spatial transcriptomics data from a single-cell spatial cohort (GSE181300) obtained from the GEO database. To deconvolute spatial transcriptomic data and infer the cell-type proportions for each spatial spot, we leveraged SPOTlight, a machine-learning-based deconvolution method. The reference profiles for deconvolution were derived from the expression patterns of 13 distinct cell types identified in our single-cell discovery cohort, ensuring a biologically meaningful representation of cellular distribution in spatial transcriptomic samples.

Identification and recovery of cell states and ecotypes using ecotyper

EcoTyper is a machine learning framework designed for the large-scale identification of cell states and multicellular communities, referred to as ecotypes. The EcoTyper pipeline consists of four key steps: (1) in silico purification, (2) cell state discovery, (3) ecotype discovery, and (4) recovery of cell states and ecotypes.

In the in silico purification step, raw gene expression data underwent rigorous preprocessing, including quality control filtering, normalization, and batch effect correction where applicable. Low-quality cells with excessive mitochondrial gene expression or low feature counts were excluded. Expression profiles were scaled to counts per million (CPM) and major cell types were identified based on Seurat clustering and annotation, encompassing B cells, plasma cells, CD4⁺ T cells, CD8⁺ T cells, NK cells, monocytes, macrophages, conventional dendritic cells (cDCs), plasmacytoid dendritic cells (pDCs), mast cells, endothelial cells, epithelial cells, and fibroblasts. To minimize potential biases introduced by cell annotation, we validated major cell type identities using canonical marker genes.

For each cell type, EcoTyper was executed 50 times using a variant of non-negative matrix factorization (NMF) to robustly define cell states. To ensure reproducibility, we set the number of clusters to range from 2 to 10 and evaluated the clustering robustness using cophenotype coefficients. The optimal cluster number was determined based on the closest approximation to a cophenotype coefficient of 0.98. To further refine cell state identification, we implemented strict filtering criteria: cell states with fewer than 10 marker genes or those classified as potential false positives based on the Adjusted False Positive Index (AFI) were excluded from downstream analyses.

In the Ecotype Discovery step, we quantified the degree of overlap between cell states using the Jaccard index, assessing statistical significance with a hypergeometric test (p<0.01). The Jaccard matrix was hierarchically clustered using average linkage and Euclidean distance, with the optimal number of ecotypes (k) selected based on silhouette width. To improve the robustness of ecotype classification in head and neck tumors, we imposed a minimum requirement of three distinct cell state components per ecotype.

In the Cell State and Ecotype Recovery step, EcoTyper applied the NMF model in a reference-based manner to restore predefined cell states from external datasets. We first ensured that the validation dataset was processed under identical quality control and normalization parameters as the discovery dataset to prevent technical biases. EcoTyper then generated a coefficient matrix from the gene expression matrix, representing each cell state as a weight. A permutation test was conducted to statistically evaluate the recovery of each cell state in the validation cohort, with statistical significance determined using a z-score threshold of 1.65.

By incorporating rigorous data preprocessing and quality control measures at each step, we ensured the robustness and reproducibility of EcoTyper-based cell state and ecotype identification in HNSC.

Assessment of the rationality of cell state clustering

To evaluate the rationality of the cell state subgroup clustering, we conducted two complementary analyses using the R packages ROGUE and scSHC. First, ROGUE was employed to assess the purity of the cell state subgroups. For each sample, the ROGUE index for every cell state was calculated and then integrated based on cell type. A bar chart displaying these values across different cell states was generated to visually illustrate the clustering purity within each cell type. In addition, the scSHC package was used to further validate the clustering results. Specifically, the testClusters function was applied to examine whether the identified subgroups were overclustered. The consistency between the clustering results from testClusters and the original clustering configuration confirmed the statistical significance and overall validity of the cell state subgroup clustering.

ST visualization and colocalization analysis

The abundance of cell states and ecotypes within each spatial barcode spot was estimated by inferring the relative abundance of HNSC cell states in the Visium array using EcoTyper. For each spot, the most abundant cell state of each cell type was normalized to 1, with all other states set to 0. To standardize ecotype abundance, the 99th percentile of all ecotype values was normalized to 1.

Colocalization of cell states was evaluated by calculating the Spearman correlation coefficient matrix for cell state abundance within each spot. Additionally, spatial clustering of ecotypes was analyzed using Moran's I statistic [27], which assessed the relative abundance of SEs in each spot in relation to their immediate neighboring spots.

Genomic mutation analysis

TCGA mutation data for HNSC were used to select the 18 most frequently mutated genes. Tumor barcodes were standardized and mutation data reshaped into a geneby-sample matrix (with multiple mutation types consolidated as "Multi_Hit"), followed by integration of clinical and ecotype data. An OncoPrint generated via the ComplexHeatmap package visualized the mutational landscapes of the E7 and E8 ecotypes.

Cell-cell interaction network

To explore potential intercellular crosstalk within each ecosystem, we employed the R package CellChat following its standard pipeline to infer the distribution and expression of ligand-receptor pairs. For the cell–cell communication analysis, we utilized all components of the CellChatDB except for the "Non-protein Signaling" category. Additionally, any interactions involving fewer than 10 cells were filtered out.

Treatment sensitivity analysis

For immunotherapy sensitivity analysis, we collected three established immunotherapy sensitivity signatures (IFN- γ 6, T cell-inflamed GEP, and TLS [23–25]) along with immune-related markers (CD8A, MS4A1, FOXP3, CD274, CD163) and exhaustion/activation-related markers (PDCD1, CTLA4, IFNG) previously used to assess the utility of Immunoscore in melanoma [26]. These markers were log2-transformed and scaled to unit variance across pre-treatment samples in each dataset.

All immunotherapy and chemotherapy gene expression datasets were TPM-normalized prior to analysis, and only RNA-seq profiles from pre-treatment tumors were included. To mitigate potential batch effects, each indicator was independently estimated within each dataset. We applied univariate Cox proportional hazards regression to evaluate each indicator's association with overall survival, extracting the z-score for each. Additionally, the binary association between each indicator and treatment response was assessed using a two-sided Wilcoxon test, with z-scores calculated from the Wilcoxon p-values. Finally, the resulting z-score rankings were averaged across outcome associations and treatment types to generate a final ranking for each indicator.

CE network visualization

The igraph package was used to construct a weighted, undirected network representing the relationships

between cell states within each CE. The edge weights were determined by the Jaccard index between cell states, reflecting the degree of overlap between them. The network layout was generated using the layout_ with_fr function, which positions the nodes based on the Fruchterman-Reingold force-directed algorithm.

Statistics and reproducibility

To evaluate the clinical relevance of cell states and ecotypes in HNSC, we performed survival analyses using the 'survfit' function from the 'survival' package. Cutoff points for high and low abundance were determined with the 'surv_cutpoint' function from the 'survminer' package, and statistical significance was assessed using two-sided log-rank tests. Univariate Cox proportional hazards models were employed to identify risk and protective cell states and ecotypes, with the c-index values calculated via the 'coxph' function in the 'survival' package. The Wilcoxon rank-sum test was applied to compare the distributions of unpaired and paired samples. Correlations between variables were analyzed using the Spearman correlation coefficient. Statistical significance was defined as a p-value less than 0.05. All statistical analyses were conducted using Prism (GraphPad Software) or R within the RStudio environment.

Discussion

In this study, we applied the EcoTyper machine learning framework to identify and characterize cell states and multicellular communities (ecotypes) in HNSC. Unlike previous pan-cancer classifications of tumor cell states and ecotypes, our approach enables the identification of HNSC-specific cell states and ecotypes [13, 20]. By analyzing the various cell states within the tumor microenvironment and their community structures, we provided a comprehensive characterization of the heterogeneity within the HNSC tumor microenvironment. Analyses across independent datasets revealed significant associations between specific cell states, tumor ecotypes, and patient outcomes, underscoring the critical role of tumor microenvironment heterogeneity in HNSC prognosis and treatment response. This study offers a novel perspective for precision medicine in HNSC, extending beyond conventional TNM staging.

To systematically characterize the HNSC tumor microenvironment, we integrated primary HNSC single-cell sequencing data from multiple cohorts to create a discovery cohort for the ecotype framework [28]. We further validated the identified cell states and ecotypes using metastatic HNSC single-cell sequencing data, spatial transcriptomics samples, and single-cell sequencing data from other cancer types. The results showed that most cell states and ecotypes were reproducible and were applicable to other cancers, demonstrating the robustness of the ecotype-based classification.

Among the identified cell states, several "key" states were strongly associated with patient prognosis [29–33]. Notably, across multiple cohorts, a high abundance of SPP1+M2-like macrophages and ACTA2+fibroblast emerged as the most significant factors associated with poor prognosis in HNSC patients. We further validated the association between ACTA2+fibroblast abundance and prognosis in our own TMA cohort, confirming that a higher abundance of ACTA2+fibroblast correlated with worse outcomes. These findings suggest that the abundance of fibroblast in the tumor microenvironment may be a critical determinant of HNSC prognosis.

Conversely, patients with a higher abundance of ATPIF1+ATP5G1+plasma cell, which are related to mitochondrial metabolism, has significantly better outcomes. ATPIF1 and ATP5G1 play key roles in mitochondrial function-ATP5G1, a component of the ATP synthase complex, drives ATP production [34], while ATPIF1 protects mitochondrial function under hypoxic conditions by inhibiting ATP synthase activity [35]. The synergistic effects of these proteins may enhance mitochondrial function in plasma cells, promoting their survival and immune activity within the tumor microenvironment [36], ultimately leading to a stronger immune response and improved patient prognosis.

Since each tumor ecotype integrates multiple cell state contributions, we hypothesize that ecotype analysis could enhance clinical outcome prediction. Indeed, we observed a strong correlation between tumor ecotypes and patient prognosis. For instance, the E7 ecotype, predominantly composed of classical dendritic cells (cDC1) and natural killer T (NKT) cells, was associated with favorable outcomes. Our analysis demonstrated that the prognostic accuracy of the E7 ecotype was comparable to, and in some instances surpassed, that of existing immunotherapy predictors.

Compared to previous pan-cancer ecotype classifications, our HNSC-specific ecotypes demonstrated superior prognostic and predictive capabilities. Unlike general classifications that broadly group tumors based on molecular and cellular features, our ecotype framework captures tumor-specific cellular states and ecological niches in HNSC. This fine-grained classification allows for a more precise stratification of HNSC based on cellular compositions, providing better prognostic insights. Notably, E7 and E8 ecotypes outperformed CE1 and CE10 in predicting patient outcomes across multiple cohorts (Fig. 5E), highlighting the enhanced accuracy of our classification.

These findings suggest that detailed characterization of the tumor microenvironment based solely on ecotype features could provide accurate predictions of patient responses to immunotherapy, underscoring the potential of the ecotype framework in precision oncology.

In contrast, the E8 ecotype is primarily composed of TGFBI+fibroblast and TGFBI+macrophages, and is strongly associated with poor prognosis. TGFBI plays a pivotal role in the tumor microenvironment, particularly by activating TGFBI + fibroblast and remodeling the extracellular matrix, thereby facilitating tumor cell invasion and migration [37-39]. Additionally, TGFBI+macrophages contribute to immunosuppression by creating a microenvironment that inhibits T cell activation and function, thereby promoting tumor growth [40-42]. We observed that patients with the E8 ecotype exhibited poor responses to immunotherapy, likely due to the combined immunosuppressive effects of TGFBI+fibroblast and TGFBI+macrophages within the tumor microenvironment. The high expression of these cells may suppress T cell-mediated anti-tumor activity, thereby compromising the effectiveness of immunotherapy. This finding highlights the critical inhibitory role of TGFBI+fibroblast and TGFBI + macrophages in shaping the tumor immune microenvironment, particularly in the E8 ecotype, where their elevated expression could significantly impact the patient's immune response. Targeted therapeutic strategies aimed at TGFBI+fibroblast, TGFBI+macrophages, or their associated signaling pathways may help counteract this immunosuppressive environment, potentially improving the response to immunotherapy in these patients. These insights further underscore the value of ecotype analysis as a guiding tool for personalized treatment approaches.

In addition, we conducted an in-depth analysis of the relationship between cell states, ecotypes, and chemotherapy sensitivity in HNSC. The results revealed that multiple cell states within the E1 ecotype were significantly associated with better chemotherapy response. As a combination of cell states, the E1 ecotype outperformed the most robust predictor of favorable chemotherapy outcomes in patients.

Moreover, compared to the previous pan-cancer ecotype classification, the predictive efficacy of E1 and fibroblast (S02) for chemotherapy sensitivity was notably stronger, suggesting the superior ability of this classification in guiding treatment decisions (Fig. 6A). Unlike pan-cancer classifications, which provide broad tumor stratifications, our study demonstrates that tumor-specific ecotypes are more effective in identifying chemotherapy-sensitive and chemotherapy-resistant subgroups. This refined classification could thus facilitate personalized treatment selection, improving therapeutic outcomes for HNSC patients.

In contrast, ACTA2+fibroblast, as a distinct and independent cell state, surpassed all other individual cell states and their combinations to become a core marker of chemotherapy resistance. This cell state likely undermines chemotherapy efficacy by altering the tumor microenvironment, enhancing tumor cell survival, and limiting drug penetration [43, 44].

Notably, among all cell states and ecotypes, ACTA2+fibroblast exhibit the strongest predictive capacity for chemotherapy resistance, underscoring their pivotal role in this process. Recent studies have also shown that TSPAN8+fibroblast contribute to chemotherapy resistance in breast cancer [45]. Consistent with these findings, our study demonstrates a significant association between the abundance of fibroblast populations and chemotherapy resistance in patients with HNSC, further linking it to poor prognosis. These results highlight the potential of fibroblast as a key markers of chemotherapy resistance and adverse clinical outcomes in HNSC.

However, while our study provides valuable insights into the HNSC tumor microenvironment, there are several areas that could be further explored. Future research could integrate multi-omics data (e.g., proteomics, metabolomics) to deepen our understanding of the molecular mechanisms underlying tumor progression and immune evasion. Additionally, applying advanced machine learning techniques, such as deep learning or multi-task learning, could refine our cell state and ecotype classifications and reveal novel biomarkers for prognosis and treatment response. Lastly, incorporating longitudinal data would enable the study of dynamic changes in the tumor microenvironment over the course of treatment, offering more predictive power for patient monitoring and personalized therapy.

In summary, we provided a comprehensive characterization of the tumor microenvironment heterogeneity in HNSC by analyzing large-scale bulk and single-cell transcriptome data. Our findings offer new insights for predicting prognosis and evaluating potential treatment responses, contributing to the advancement of personalized cancer treatment.

Abbreviations

HNSC Head and neck squamous cell carcinoma

TME Tumor microenvironment

- AJCC The American Joint Committee on Cancer
- TMA Tissue microarray

- ST Spatial transcriptomics
- CE Carcinoma ecotype

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06299-4.

Supplementary material 1: Figure S1. Summary of patient cohorts and datasets used for the discovery and validation of HNSC cell states and ecosystems. This study included seven HNSC single-cell cohorts, integrating 104 primary tumor samples as the discovery cohort for identifying cell states and ecotypes, and 27 metastatic samples as the validation cohort for confirming cell states. Additionally, single-cell samples from bladder cancer and melanoma were used to extend the cell state discovery from HNSC to the pan-cancer context.Four cohorts from TCGA and GEO were utilized to explore the relationship between the abundance of HNSC cell states and ecotypes and patient prognosis. Furthermore, three immunotherapy cohorts from the TIGERand Xiangya chemotherapy cohort were analyzed to assess the correlation between cell states/ecotypes and the response to immunotherapy and chemotherapy. Spatial transcriptomic analysis of 8 single HNSC samples from the GEO database was used to verify the spatial distribution of cell states and ecotypes. Figure S2. Marker gene expression in single-cell cohorts.Gravity density map of marker gene expression for cell subpopulations in the discovery cohort. UMAP plot of the validation cohort consisting of single-cell samples from metastatic HNSC patients.Marker gene expression of 13 cell subpopulations in the validation cohort. The ROGUE index of 13 HNSC cell types integrates the ROGUE indices of different cell states for each cell type. Figure S3. Clinical association of cell state abundance. Survival associations of 66 cell states in the GEO integrated cohort, with the most significant marker genes for unfavorable and favorable cell states indicated. Comparison of c-index values for cell state abundance and clinical stage across different cohorts. Multicolor immunofluorescence of TMA, a total of 94 tissue samples with high staining guality and suitable for analysis were obtained. Layout of High-Quality Tissue MicroarraySpots. Relationship between HPV status and the abundance of cell states with favorable prognosis in the TCGA cohort. Relationship between HPV status and the abundance of cell states with unfavorable prognosis in the TCGA cohort. Figure S4. Spatial distribution of HNSC cell states and ecotypesSpatial distribution of each cell state within the E7 ecotype, analyzed using spatial transcriptomics. Spatial distribution of each cell state within the E1 ecotype, analyzed using spatial transcriptomics. Spatial distribution of tumor ecotypes across 8 HNSC samples.Network diagrams depicting putative ligand-receptor interactions between cell states within each HNSC ecotype. Figure S5. Ecotype-clinical associations. Histogram displaying c-index values for tumor ecotypes and clinical stages. Scatter plots illustrating the correlations between ecotype survival associations in the TCGA-HNSC cohort and the GEO integrated cohort.Correlation between HNSC ecotypes and clinicopathological parameters.UMAP plots of single-cell sequencing cohorts for BLCAand melanoma.Recovery of cell states in single-cell samples from BLCAand melanoma. Figure S6. ROC and survival curves for E1 ecotype and fibroblast.ROC curves for E1 ecotype and fibroblast. Survival curve for fibroblastin the TCGA-HNSC chemotherapy cohort. Survival curve for fibroblastin the Xiangya chemotherapy cohort. Table S1. Overview of gene expression cohorts analyzed in this study. Table S2 Marker genes in 66 cell states discovered by EcoTyper. Table S3. Recovery of HNSC cell states in different scRNA-Seq cohorts. Table S4. Assessment of Cluster Significance in Single-Cell RNA Sequencing Data Using the scSHC testClusters Function. Table S5. CellChat-Based Analysis of Ligand-Receptor Interactions Across HNSC Ecotypes. Table S6. Association of HNSC ecotypes and cell states abundance with survival and response to immunotherapy and chemotherapy. Table S7. Clinical information and proportion of ACTA2+COL1A1+ double-positive cells in TMA samples.

Acknowledgements

This study was supported by the Postgraduate Research Innovation Project of Hunan Province (CX20230280), the National Natural Science Foundation of China (Nos. 82473442, 82103631) and the Youth Science Foundation of

Xiangya Hospital (No. 2020Q03). This work was supported in part by the High Performance Computing Center of Central South University.

Author contributions

L. X led the study's conception, overall design, main data analysis, and paper writing. Z. S and Z. P contributed to data acquisition and analysis. Y. Q directed key data analysis efforts. D. H provided valuable advice on paper revision and data analysis. Y. L played a crucial role in paper writing and revision. C. L and X. Z made significant contributions to study initiation, supervision, project, management, manuscript writing, and final revision. The final manuscript was read and approved by all authors.

Funding

The Postgraduate Research Innovation Project of Hunan Province (CX20230280) the National Natural Science Foundation of China (Nos. 82473442, 82103631) the Youth Science Foundation of Xiangya Hospital (No. 2020Q03).

Availability of data and materials

The data and code used in this study are available in public databases or from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Ethical approval was granted by the Ethics Committee of Xiangya Hospital, Central South University (No. 202108351). Informed consent forms were signed by all patients.

Consent for publication

All participants were thoroughly informed about the study's objectives and potential implications. They voluntarily provided written informed consent, agreeing to the use of their personal data for the research and to the publication of the study's findings.

Competing interests

The authors declare that they have no competing interests in this section.

Received: 24 November 2024 Accepted: 23 February 2025 Published online: 01 March 2025

References

- Chow LQM. Head and neck cancer. N Engl J Med. 2020;382:60–72. https:// doi.org/10.1056/NEJMra1715715.
- Li Q, Tie Y, Alu A, Ma X, Shi H. Targeted therapy for head and neck cancer: signaling pathways and clinical studies. Signal Transduct Target Ther. 2023;8:31. https://doi.org/10.1038/s41392-022-01297-0.
- Fasano M, et al. Immunotherapy for head and neck cancer: present and future. Crit Rev Oncology/Hematol. 2022;174:103679. https://doi.org/10. 1016/j.critrevonc.2022.103679.
- 4. Johnson DE, et al. Head and neck squamous cell carcinoma. Nat Rev Dis Primers. 2020;6:92. https://doi.org/10.1038/s41572-020-00224-3.
- Ruffin AT, et al. Improving head and neck cancer therapies by immunomodulation of the tumour microenvironment. Nat Rev Cancer. 2023;23:173–88. https://doi.org/10.1038/s41568-022-00531-9.
- Ren S, et al. Intratumoral CD103(+) CD8(+) T cells predict response to neoadjuvant chemoimmunotherapy in advanced head and neck squamous cell carcinoma. Cancer Commun (Lond). 2023;43:1143–63. https:// doi.org/10.1002/cac2.12480.
- Lydiatt WM, et al. Head and neck cancers major changes in the American Joint Committee on cancer eighth cancer staging manual. CA Cancer J Clin. 2017;67:122–37. https://doi.org/10.3322/caac.21389.
- Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. Nat Rev Immunol. 2018;18:35–45. https://doi.org/10.1038/ nri.2017.76.

- Tang F, et al. A pan-cancer single-cell panorama of human natural killer cells. Cell. 2023;186:4235-4251.e4220. https://doi.org/10.1016/j.cell.2023. 07.034.
- Cheng S, et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell. 2021;184:792-809.e723. https://doi.org/10. 1016/j.cell.2021.01.010.
- Bill R, et al. CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. Science. 2023;381:515–24. https://doi.org/10.1126/science.ade2292.
- Liu ZL, et al. Single cell deciphering of progression trajectories of the tumor ecosystem in head and neck cancer. Nat Commun. 2024;15:2595. https://doi.org/10.1038/s41467-024-46912-6.
- Luca BA, et al. Atlas of clinically distinct cell states and ecosystems across human solid tumors. Cell. 2021;184:5482-5496.e5428. https://doi.org/10. 1016/j.cell.2021.09.014.
- 14. Quah HS, et al. Single cell analysis in head and neck cancer reveals potential immune evasion mechanisms during early metastasis. Nat Commun. 2023;14:1680. https://doi.org/10.1038/s41467-023-37379-y.
- Puram SV, et al. Cellular states are coupled to genomic and viral heterogeneity in HPV-related oropharyngeal carcinoma. Nat Genet. 2023;55:640–50. https://doi.org/10.1038/s41588-023-01357-3.
- Chu Y, et al. Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. Nat Med. 2023;29:1550–62. https://doi.org/ 10.1038/s41591-023-02371-y.
- Choi JH, et al. Single-cell transcriptome profiling of the stepwise progression of head and neck cancer. Nat Commun. 2023;14:1055. https://doi. org/10.1038/s41467-023-36691-x.
- Yang Y, et al. Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes. Cell. 2024;187:4790-4811.e4722. https://doi.org/ 10.1016/j.cell.2024.06.038.
- Yuan W, et al. Dual role of CXCL8 in maintaining the mesenchymal state of glioblastoma stem cells and M2-like tumor-associated macrophages. Clin Cancer Res. 2023;29:3779–92. https://doi.org/10.1158/1078-0432. Ccr-22-3273.
- Steen CB, et al. The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. Cancer Cell. 2021;39:1422-1437.e1410. https://doi.org/10.1016/j.ccell.2021.08.011.
- Subramanian A, et al. Sarcoma microenvironment cell states and ecosystems are associated with prognosis and predict response to immunotherapy. Nat Cancer. 2024;5:642–58. https://doi.org/10.1038/ s43018-024-00743-y.
- 22. Qi J, et al. Single-cell and spatial analysis reveal interaction of FAP(+) fibroblasts and SPP1(+) macrophages in colorectal cancer. Nat Commun. 2022;13:1742. https://doi.org/10.1038/s41467-022-29366-6.
- Cabrita R, et al. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. Nature. 2020;577:561–5. https://doi.org/10.1038/ s41586-019-1914-8.
- Ayers M, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017;127:2930–40. https://doi.org/10.1172/ jci91190.
- van den Ende T, et al. Neoadjuvant chemoradiotherapy combined with atezolizumab for resectable esophageal adenocarcinoma: a single-arm phase II feasibility trial (PERFECT). Clin Cancer Res. 2021;27:3351–9. https://doi.org/10.1158/1078-0432.Ccr-20-4443.
- Holder AM, et al. Defining clinically useful biomarkers of immune checkpoint inhibitors in solid tumours. Nat Rev Cancer. 2024;24:498–512. https://doi.org/10.1038/s41568-024-00705-7.
- Moran PA. Notes on continuous stochastic phenomena. Biometrika. 1950;37:17–23.
- Storrs EP, et al. High-dimensional deconstruction of pancreatic cancer identifies tumor microenvironmental and developmental stemness features that predict survival. NPJ Precis Oncol. 2023;7:105. https://doi. org/10.1038/s41698-023-00455-z.
- Liu Y, et al. Identification of a tumour immune barrier in the HCC microenvironment that determines the efficacy of immunotherapy. J Hepatol. 2023;78:770–82. https://doi.org/10.1016/j.jhep.2023.01.011.
- Matusiak M, et al. Spatially segregated macrophage populations predict distinct outcomes in colon cancer. Cancer Discov. 2024;14:1418–39. https://doi.org/10.1158/2159-8290.Cd-23-1300.

- 31. Lavie D, Ben-Shmuel A, Erez N, Scherz-Shouval R. Cancer-associated fibroblasts in the single-cell era. Nat Cancer. 2022;3:793–807. https://doi.org/10.1038/s43018-022-00411-z.
- Luo H, et al. Pan-cancer single-cell analysis reveals the heterogeneity and plasticity of cancer-associated fibroblasts in the tumor microenvironment. Nat Commun. 2022;13:6619. https://doi.org/10.1038/ s41467-022-34395-2.
- Ma C, et al. Pan-cancer spatially resolved single-cell analysis reveals the crosstalk between cancer-associated fibroblasts and tumor microenvironment. Mol Cancer. 2023;22:170. https://doi.org/10.1186/ s12943-023-01876-x.
- Sun L, et al. RNA-binding protein RALY reprogrammes mitochondrial metabolism via mediating miRNA processing in colorectal cancer. Gut. 2021;70:1698–712. https://doi.org/10.1136/gutjnl-2020-320652.
- Acin-Perez R, et al. Inhibition of ATP synthase reverse activity restores energy homeostasis in mitochondrial pathologies. Embo J. 2023;42:111699. https://doi.org/10.15252/embj.2022111699.
- Zhong G, et al. scRNA-seq reveals ATPIF1 activity in control of T cell antitumor activity. Oncoimmunology. 2022;11:2114740. https://doi.org/ 10.1080/2162402x.2022.2114740.
- Kisoda S, et al. Prognostic value of partial EMT-related genes in head and neck squamous cell carcinoma by a bioinformatic analysis. Oral Dis. 2020;26:1149–56. https://doi.org/10.1111/odi.13351.
- Puram SV, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. Cell. 2017;171:1611-1624.e1624. https://doi.org/10.1016/j.cell.2017.10.044.
- Yang R, et al. Downregulation of nc886 contributes to prostate cancer cell invasion and TGFβ1-induced EMT. Genes Dis. 2022;9:1086–98. https://doi. org/10.1016/j.gendis.2020.12.010.
- Lecker LSM, et al. TGFBI production by macrophages contributes to an immunosuppressive microenvironment in ovarian cancer. Cancer Res. 2021;81:5706–19. https://doi.org/10.1158/0008-5472.Can-21-0536.
- Zhou J, et al. A novel role of TGFBI in macrophage polarization and macrophage-induced pancreatic cancer growth and therapeutic resistance. Cancer Lett. 2023;578: 216457. https://doi.org/10.1016/j.canlet. 2023.216457.
- Huang H, Tang Q, Li S, Qin Y, Zhu G. TGFBI: a novel therapeutic target for cancer. Int Immunopharmacol. 2024;134: 112180. https://doi.org/10. 1016/j.intimp.2024.112180.
- Lodyga M, Hinz B. TGF-β1—A truly transforming growth factor in fibrosis and immunity. Semin Cell Dev Biol. 2020;101:123–39. https://doi.org/10. 1016/j.semcdb.2019.12.010.
- Cremasco V, et al. FAP delineates heterogeneous and functionally divergent stromal cells in immune-excluded breast tumors. Cancer Immunol Res. 2018;6:1472–85. https://doi.org/10.1158/2326-6066.Cir-18-0098.
- Fan G, et al. TSPAN8(+) fibroblastic cancer-associated fibroblasts promote chemoresistance in patients with breast cancer. Sci Transl Med. 2024;16:eadj5705. https://doi.org/10.1126/scitranslmed.adj5705.

Publisher's Note

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