



Integrated analysis of single-cell RNA-seq and bulk RNA-seq reveals distinct cancer-associated fibroblasts in head and neck squamous cell carcinoma

Qian Zhang^{1#}, Yuxin Wang^{1#}, Chengwan Xia¹, Liang Ding², Yumei Pu¹, Xiaobei Hu³, Huiming Cai⁴, Qingang Hu¹

¹Department of Oral and Maxillofacial Surgery, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China; ²Central Laboratory of Stomatology, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China; ³Department of Orthodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China; ⁴Nanjing Nuoyuan Medical Devices Co., Ltd., Nanjing, China

Contributions: (I) Conception and design: Q Zhang; (II) Administrative support: H Cai; (III) Provision of study materials or patients: Q Hu; (IV) Collection and assembly of data: Y Wang; (V) Data analysis and interpretation: Q Zhang, L Ding; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Huiming Cai. Nanjing Nuoyuan Medical Devices Co., Ltd., Nanjing, China. Email: chm9166@aliyun.com; Qingang Hu. Department of Oral and Maxillofacial Surgery, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China. Email: qghu@nju.edu.cn.

Background: The heterogeneity of cancer-associated fibroblasts (CAFs) in head and neck squamous cell carcinoma (HNSCC) has been widely acknowledged, but has not yet been elucidated. The potential roles and clinical relevance of CAFs subclusters in HNSCC progression remain obscure.

Methods: In this study, we combined single-cell and bulk tissue transcriptome profiles of HNSCC with clinical data from The Cancer Genome Atlas (TCGA). The *Seurat* package was used to perform single-cell RNA-seq analysis to distinguish distinct CAFs subtypes. Prognostic relevance of several CAFs markers was assessed and functional analysis was also performed.

Results: We identified eight CAFs subclusters; of these, seven showed enhanced expression levels in HNSCC tumor tissues compared to normal tissue, and three (clusters 0, 3, and 4) were associated with poorer overall survival. Further functional analysis revealed that cluster 0 was characterized by myofibroblasts with high alpha smooth muscle actin (aSMA) expression and enrichment in smooth muscle contraction. The cluster 3 exhibited expression of extracellular matrix (ECM)-related genes and was enriched in epithelial-mesenchymal transition (EMT)-related gene sets. Cluster 4 expressed high levels of the major histocompatibility complex (MHC) class II family, which was characterized as antigen-presenting CAFs.

Conclusions: We determined CAFs heterogeneity in HNSCC. 8 CAFs subclusters were recognized and 3 of which were prognosis related. The 3 CAFs subclusters showed distinct phenotypes enriched in myofibroblast function, ECM remodeling and antigen-presenting function respectively.

Keywords: Head and neck squamous cell carcinoma (HNSCC); tumour heterogeneity; cancer-associated fibroblasts (CAFs); prognosis

Submitted May 07, 2021. Accepted for publication Jun 15, 2021.

doi: 10.21037/atm-21-2767

View this article at: <https://dx.doi.org/10.21037/atm-21-2767>

Introduction

The tumor microenvironment (TME) is a complex ecosystem composed of cancer cells, immune cells, and stromal cells (1). Cancer-associated fibroblasts (CAFs) are a major component of the TME, and their prominent roles in head and neck squamous cell carcinoma (HNSCC) progression have been widely studied (2). CAFs help HNSCC growth by secreting proteins, miRNAs and metabolites (3). On one side, they secrete factors that act on tumor cells resulting to more aggressive phenotype. On the other side, they contribute to TME reprogramming which favors tumor progress (4). Presently, the origin, phenotype, and functions of CAFs are poorly understood. They are broadly defined as spindle-shaped cell clusters, which are negative for epithelial, endothelial, and leukocyte markers (5). Numerous studies have suggested that CAFs can originate from a variety of cells, such as resident fibroblasts, precursor cells, or even tumor cells (6). Multiple origins decide the complex phenotypes and functions of CAFs in the TME. Particularly, the inter- and intra-tumor heterogeneity of CAFs has been largely demonstrated, which requires specific research on distinct tumor types at the single cell level.

A major challenge in distinguishing subclusters of CAFs is the lack of reliable markers. Development of single-cell sequencing and related algorithms provides powerful methods for the study of CAFs subclusters. Different from traditional bulk RNA sequencing, single-cell sequencing technology obtains transcriptome profiles at the level of a single-cell which makes single-cell heterogeneity studies possible (7). Li *et al.* (8) revealed two subclusters of CAFs, one of which expressed genes related to extracellular matrix (ECM) remodeling and the other expressed markers of myofibroblasts. Bartoschek *et al.* recognized four clusters CAFs in breast cancer, which were named vascular CAF, matrix CAF, cycling CAF, and developmental CAF (9). Sebastian *et al.* (10) studied CAFs heterogeneity in 4T1 breast cancer model and identified myofibroblastic CAFs, inflammatory CAFs and another subpopulation expressing major histocompatibility complex. Zhang *et al.* (11) identified 6 distinct fibroblast subsets in human intrahepatic cholangiocarcinoma, of which the majority were vascular CAFs with positive CD146 expression.

The extent of CAF heterogeneity and functions of CAFs subcluster in HNSCC remains poorly characterized. Puram *et al.* performed transcriptomes of approximately 6,000 single cells from 18 HNSCC patients using Smart-Seq

technology. The ~1,500 fibroblasts were subdivided into two main subsets and one minor subset. Typical markers, such as alpha smooth muscle actin (aSMA/ACTA2), fibroblast activation protein (FAP), and podoplanin (PDPN), were applied to distinguish different CAFs subclusters. However, functions of CAFs subclusters and their corresponding roles in HNSCC remain obscure.

In this study, we accessed HNSCC single-cell RNA-seq data and bulk RNA-seq data from a public database. Re-analysis was performed using the *Seurat* package (12) to distinguish CAFs subclusters and identify cluster biomarkers. Based on an integrated analysis from individual CAFs, bulk RNA-seq data, and clinical data, we further determined the functions of several CAF clusters and described their relationships with clinical outcomes. We present the following article in accordance with the MDAR reporting checklist (available at <https://dx.doi.org/10.21037/atm-21-2767>).

Methods

Data processing

Single-cell transcriptome files of HNSCC tumors (GSE103322) were downloaded from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>), and the data of 1,423 fibroblasts were extracted for further analysis. The Cancer Genome Atlas (TCGA) bulk RNA-sequencing (RNA-seq) data and clinical information regarding HNSCC were downloaded via cBioPortal. R software (version 3.5.3) was used for all analyses. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Single-cell RNA-seq analysis

The *Seurat* package was used to perform single-cell RNA-seq analysis. Cells that had either over 2,500 or below 200 expressed genes were removed. Also, cells that expressed over 5% of genes derived from the mitochondrial genome had been also removed due to a risk of low quality. A total of the top 2,000 variably expressed genes were selected after log normalization. Non-linear dimensional reduction was performed using the Uniform Manifold Approximation and Projection (UMAP) method, with 15 principal components and resolution at 0.7. All CAFs were then divided into 8 subclusters automatically using *FindCluster* function. Cluster biomarkers were found using the *Seurat* package. The

raw data of marker gene expression in 8 CAFs subclusters analyzed in this study was also uploaded as supplemental information in total online: <https://cdn.amegroups.cn/static/public/atm-21-2767-1.xlsx>.

Correlation to bulk RNA-seq in TCGA data

To explore potential roles of CAFs subclusters in HNSCC, the top 100 marker genes for each cluster were selected as representation. The HNSCC bulk RNA-seq data were obtained from TCGA after z-score normalization. The corresponding expression of CAFs subclusters marker genes were combined in HNSCC bulk RNA-seq data, and the average expression for each subcluster was compared between the tumor and normal groups.

Survival analysis

To assess the correlation between each CAFs subcluster and the prognosis of HNSCC patients, clinical data were downloaded using TCGA biolinks. The average combined expression of the top 100 marker genes for each subcluster was calculated for each patient. After matching the expression of selected marker genes in the bulk RNA-seq data from TCGA, 439 HNSCC samples with complete expression profiles and clinical data were included. We then assessed the effect of marker gene expression for each subcluster on overall survival using the *survival* package (<https://rdocumentation.org/packages/survival/versions/2.42-3>).

Functional enrichment analysis

To explore the potential function of each CAFs subcluster, single-sample GSEA (ssGSEA) analyses for single-cell sequencing data were performed based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and hallmark gene sets using *GSEA* package (13). All analyzed gene sets were obtained from the Molecular Signatures Database (MSigDB) (14). $P < 0.05$ was set as the cutoff.

Statistical analysis

Statistical analyses were performed using the R software (<https://mirrors.tuna.tsinghua.edu.cn/CRAN/>). The unpaired *t*-test was used to compare the marker gene expression levels of each cluster between tumor and normal

tissues. Kaplan-Meier survival analysis was conducted to compare the prognostic risk HNSCC patients with different average expression levels of each CAFs subcluster. Differences with a P value < 0.05 were considered significant.

This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Results

Single-cell analysis revealed fibroblasts heterogeneity in HNSCC

In our sample, 1,423 fibroblasts data were included. The single-cell RNA-seq data was analyzed using the *Seurat* package. After removing low quality cells, the top 2,000 variably expressed genes were included to identify CAFs subclusters (*Figure 1A*). The top 10 variably expressed genes were *ACTA2*, *HP*, *CCL20*, *SLN*, *CCL21*, *DCN*, *LUM*, *IBSP*, *DLK1*, and *TAGLN*. Subclustering revealed eight distinct subclusters (*Figure 1B,C*).

Marker genes for CAFs subclusters

After dimensionality reduction using principle component analysis, we recognized eight CAFs subclusters. To further identify the marker genes for each of these eight subclusters, we used the *FindMarker* function of the *Seurat* package to compare cells from each subcluster. Marker genes were required with a default threshold of $\text{LogFC} > 0.25$ and were also detectable for more than 15% of cells in this subcluster. A total of 5,551 marker genes were then calculated in total online: <https://cdn.amegroups.cn/static/public/atm-21-2767-1.xlsx>. The top three marker genes analyzed for each subcluster are displayed in *Figure 2*.

We then screened the expression of typical CAFs markers among eight CAFs, including α -SMA (*ACTA2*), FAP, FSP1 (*S100A4*), CTGF, VIM, PDGFRB, PDPN, TNC, and ITGB1 (*Figure 3*). FSP1, CTGF, VIM, and ITGB1 showed a high average expression among the eight subclusters. Most of the CAFs subclusters showed a high PDGFRB expression level, except for cluster 5. Apart from clusters 0 and 2, the remaining six CAFs subclusters all exhibited high PDPN expression levels. TNC was highly expressed in clusters 0, 3, and 4. The α SMA expression was elevated in clusters 0 and 2, while FAP was highly expressed in clusters 1, 3, and 6.

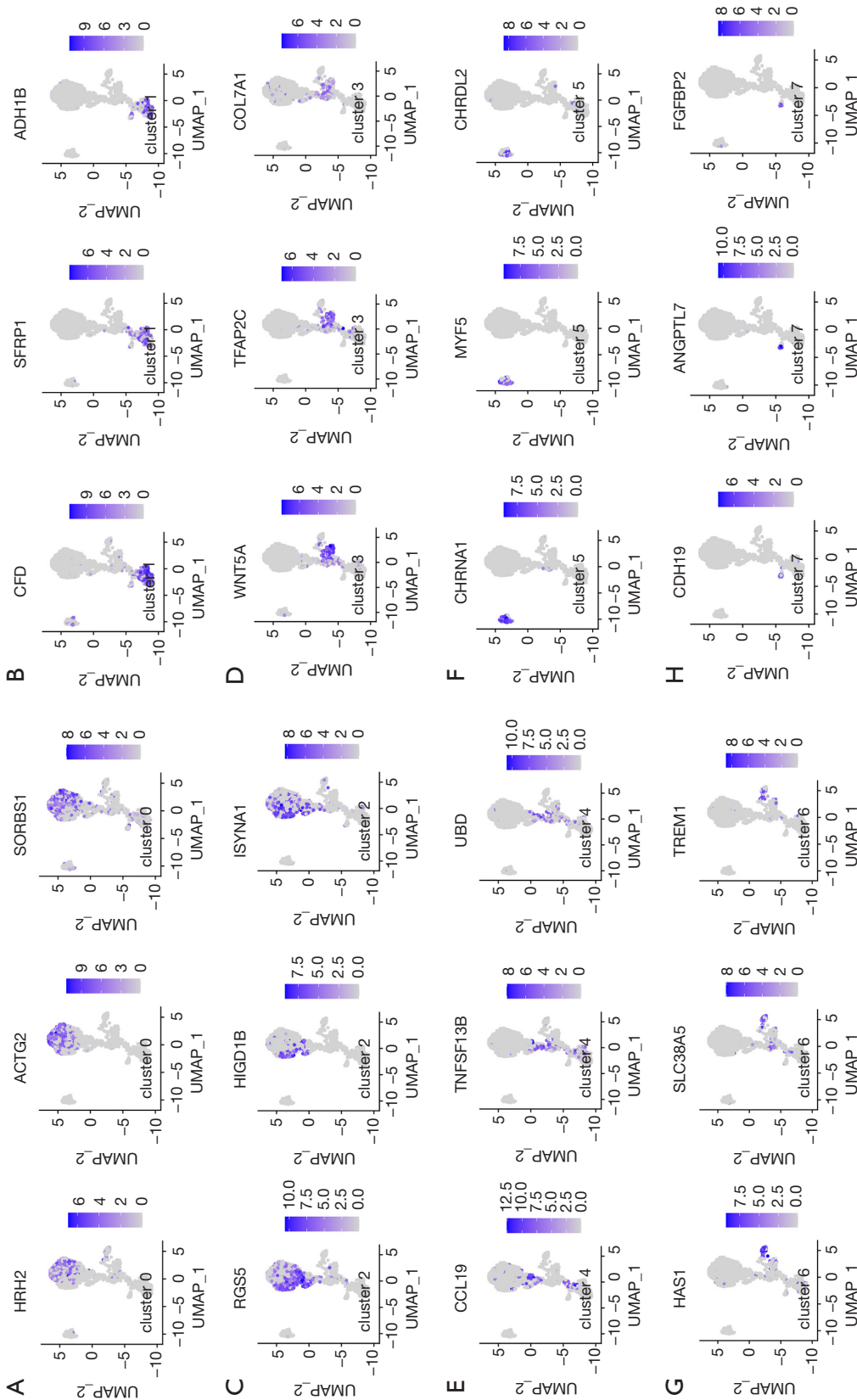


Figure 2 Top three marker genes for each cluster.

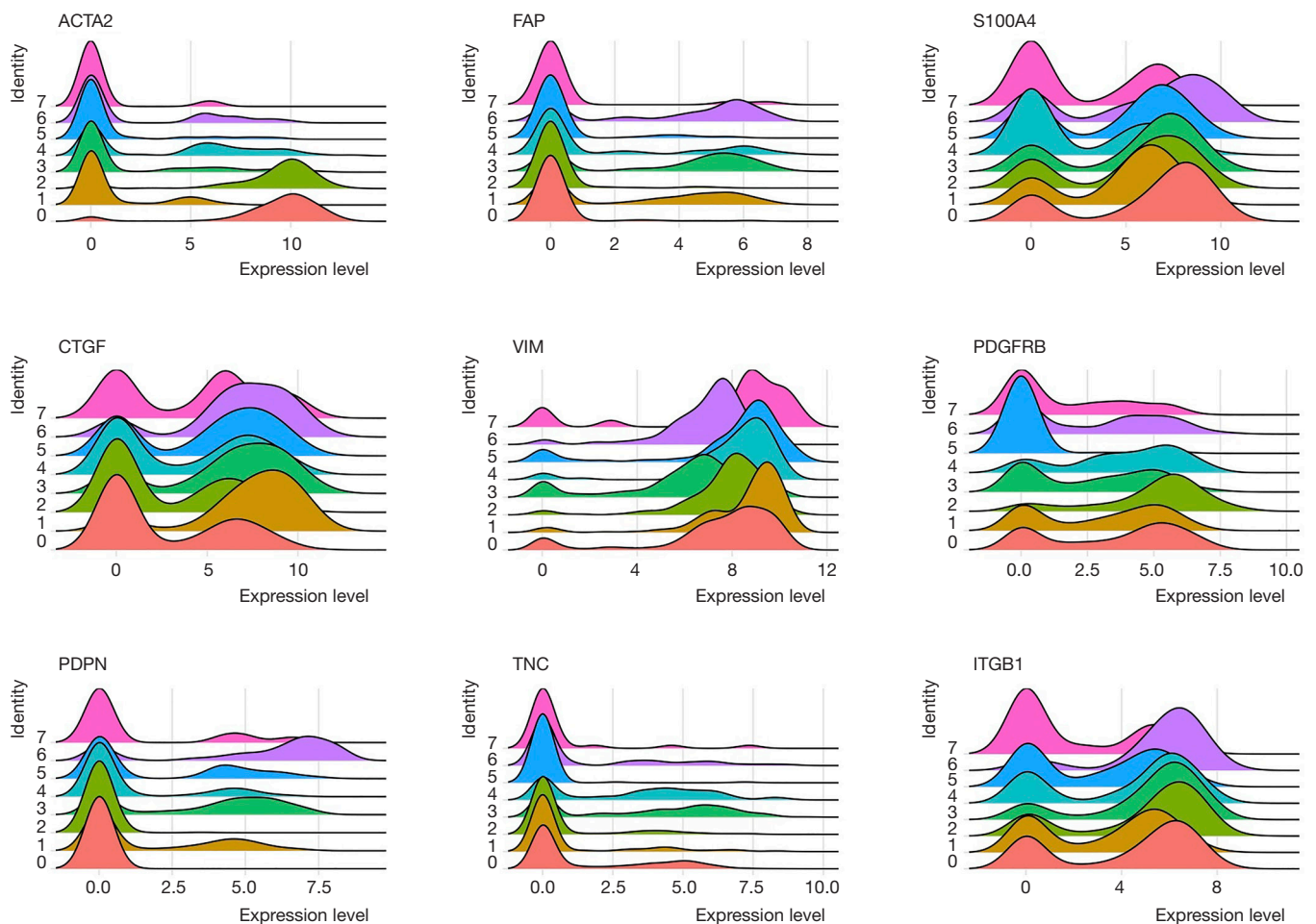


Figure 3 Ridge plot displaying typical fibroblast markers expression in eight clusters.

KEGG pathway and hallmark gene sets using ssGSEA analysis. As shown in *Figure 5A,B*, cluster 0 CAFs were enriched in the KEGG pathways of vascular smooth muscle contraction, the cyclic guanosine monophosphate-dependent protein kinase (cGMP-PKG) signaling pathway, and the oxytocin signaling pathway. The cluster 3 CAFs were enriched in the KEGG pathways of protein processing in the endoplasmic reticulum, oxidative phosphorylation, proteoglycans in cancer, and ECM-receptor interaction (*Figure 5C*). Furthermore, hallmark gene sets related to EMT were also enriched in cluster 3 CAFs (*Figure 5D*). Cluster 4 CAFs were enriched in the KEGG pathways of ribosome, antigen processing and presentation, and phagosome (*Figure 5E,F*).

Discussion

As a main component in the TME, CAFs have been reported to play multiple roles in tumor progression, including tumorigenesis, angiogenesis, metastasis, immunomodulation, metabolic reprogramming, drug resistance, and anti-tumorigenesis (15). However, no single CAFs cluster could exert all of the above functions in the TME because of the remarkable heterogeneity in their origins, phenotypes, functions, and tumor types (16).

CAF also exhibit intra-tumor heterogeneity in HNSCC. Recent studies have suggested that CAFs subclusters with high α -SMA expression exhibit greater tumor supportive ability, secreting more proteins such as matrix metalloproteinase-2 (MMP2), transforming growth factor- β

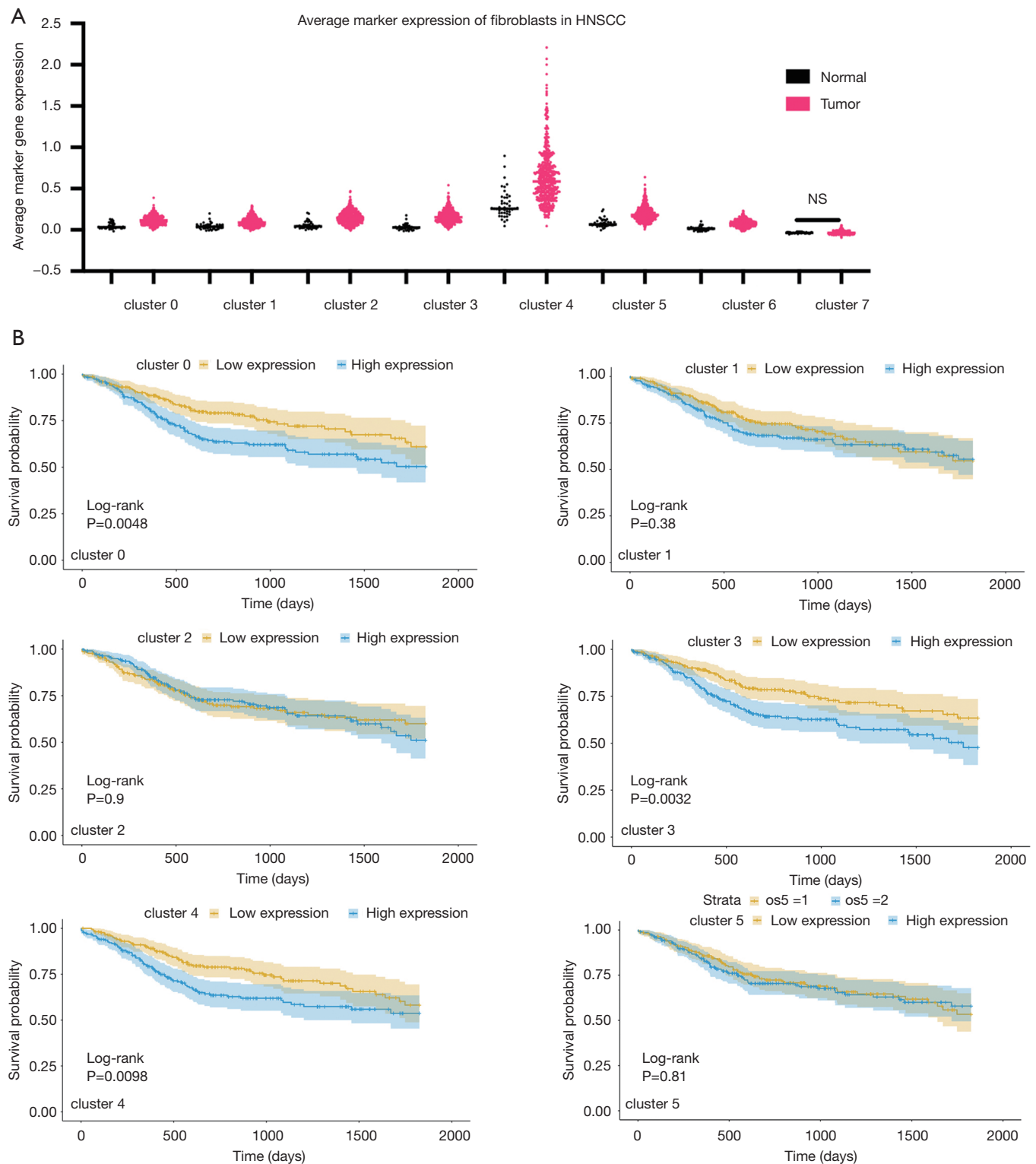


Figure 4 Cancer-associated fibroblasts (CAFs) subclusters are associated with poor survival in head and neck squamous cell carcinoma (HNSCC) patients. (A) Most CAFs subclusters (except from cluster 7) showed significantly higher expression in tumor tissues from TCGA database; (B) Kaplan-Meier curve showing the impact of the relative presence of these CAFs subclusters on HNSCC patients' overall survival. Patients with a relatively higher expression of CAFs clusters 0, 3, and 4 had a poorer overall survival ($P < 0.05$).

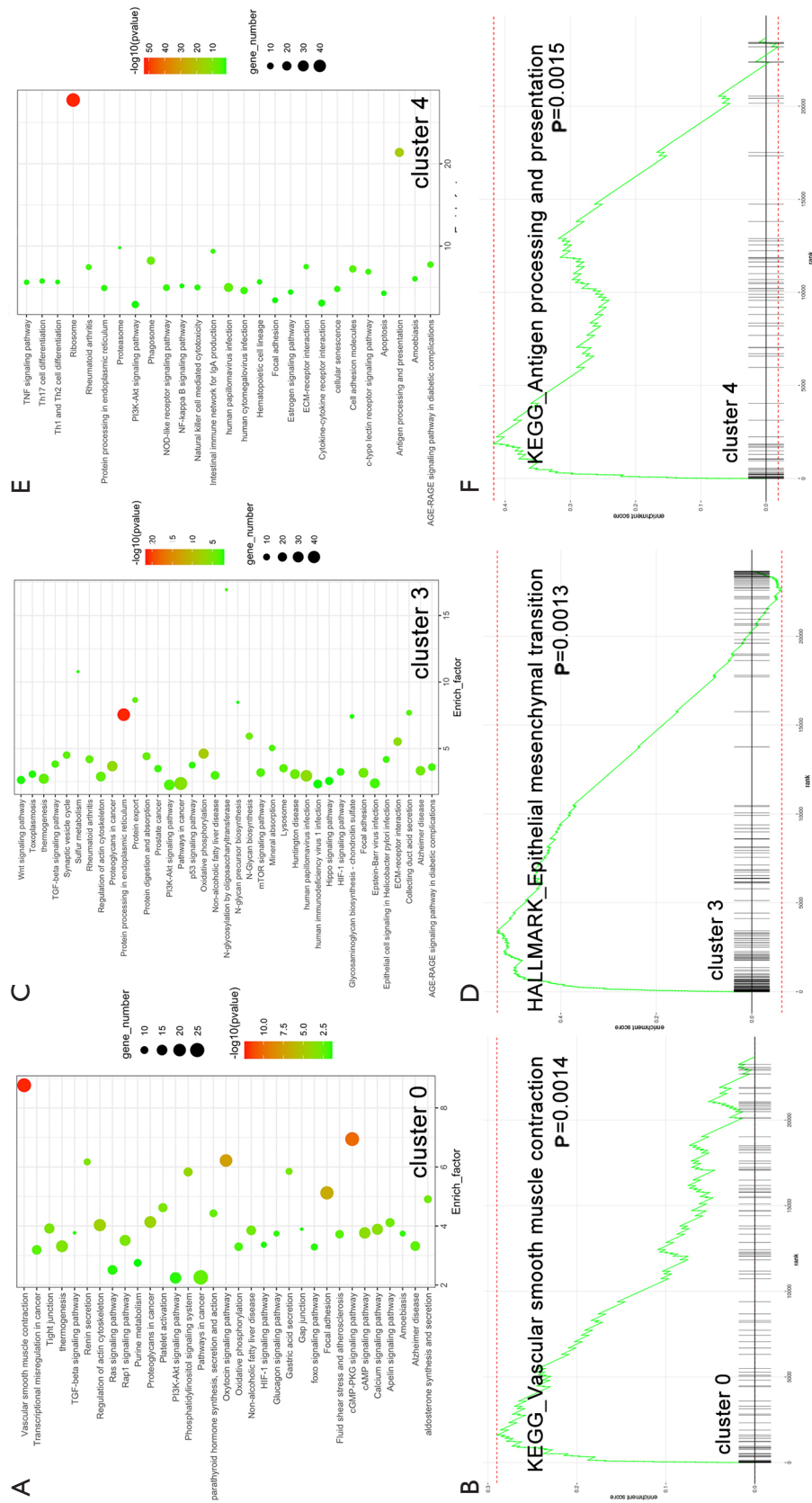


Figure 5 Functional enrichment analysis for cancer-associated fibroblasts (CAFs) subclusters. (A) Bubble chart of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for cluster 0 CAFs; (B) the single-sample gene set enrichment analysis (ssGSEA) of cluster 0 CAFs; (C) Bubble chart of KEGG pathway analysis for cluster 3 CAFs; (D) The ssGSEA of cluster 3 CAFs; (E) Bubble chart of KEGG pathway analysis for cluster 4 CAFs; (F) the ssGSEA of cluster 4 CAFs.

1/2 (TGF- β 1/2), and interleukin-6 (IL-6) (3). On the contrary, CAFs subclusters with a lack of α -SMA expression have been reported to be genetically closer to normal fibroblasts, which secrete less TGF-B1 but more hyaluronan (17). Puram *et al.* (18) profiled the transcriptomes of ~6,000 single cells from 18 HNSCC patients using SMART-Seq2 technology. Among 1,500 fibroblasts, two main subclusters and one minor subcluster were proposed. The first main subcluster was considered as myofibroblast with high α -SMA expression. The second subcluster was reported to express CAFs-related genes such as FAP, PDPN, and CTGF. The minor subcluster was considered as resting fibroblast due to a lack of typical CAFs markers.

To further illuminate the heterogeneity of CAFs in HNSCC, we applied the *Seurat* package to re-analyze the transcriptomes data of 1,423 fibroblasts downloaded from the GEO database (GSE103322). In this study, a total of eight subclusters were discriminated and corresponding marker genes were calculated. We then screened the gene expression of typical CAFs markers and found that no single marker can distinguish specific CAFs subclusters. CTGF, VIM, and ITGB1 expressed analogously high expression among all subclusters. Also, CAFs clusters 0 and 2 both exhibited high α -SMA expression, and FAP was elevated in clusters 1, 3, 4, and 6. Beyond this, recent studies also pointed out that these markers lack specificity as they are also expressed by mesodermal cells, myeloid cells, and pericytes (16). We calculated the specific marker genes for each subcluster using the *Seurat* package, which showed better discrimination (Figure 3). Further validation should be performed by *in vivo* examinations.

To explore the clinical roles of different CAFs subclusters in HNSCC, we performed integrated analysis using bulk RNA-seq data and clinical data from TCGA database. Except for cluster 7, all of the other CAFs subclusters exhibited a higher average expression level in HNSCC tumor tissues, which suggested that cluster 7 CAFs might be resting and spectators in the development of HNSCC. We further explored the prognostic risk of distinct CAFs subclusters on the overall survival of HNSCC patients. We found that HNSCC patients with high expression levels of CAFs subclusters 0, 3 and 4 had poorer overall survival, which indicated potential pro-tumor functions of these three subclusters.

SsGSEA analysis was further performed to study the functions of distinct CAFs subclusters based on the KEGG pathway and hallmark gene sets. We found that cluster 0 CAFs with high α -SMA expression were enriched in the

vascular smooth muscle contraction pathway, which was consistent with the myofibroblast function reported in previous studies (5). Myofibroblasts have been characterized by high expression of α SMA and myosin light-chain proteins (MYLK, MYL9), which were widely found in several cancers (3,19,20). Compared to resting fibroblasts, myofibroblasts exhibit enhanced proliferative and migratory properties. Apart from their well-established smooth muscle contraction ability, they have also been reported to produce collagen and TGF-B, and participate in ECM remodeling (16,21). Our results revealed a consistent smooth muscle contraction function in cluster 0, but the ECM remodeling function was more enriched in cluster 3, which had low levels of α SMA expression. This indicated that the phenotype and function of myofibroblasts should be further subdivided by speciality.

The cluster 3 CAFs were enriched in EMT gene sets. The KEGG pathways related to ECM interaction and actin cytoskeleton regulation were also enriched. Previous studies have demonstrated that CAFs are a major source of ECM components in the TME, including collagens, glycoprotein, and the matrix metalloproteinases family (MMPs) (5,22,23). Due to a remarkable ECM remodeling ability, CAFs help to establish a collagen-rich TME, which may enhance EMT and the invasiveness of cancer cells in several tumors, including breast, prostate, and colon cancers (24-26). The cluster 3 CAFs exhibited a similar phenotype and function with matrix CAF, which was characterized by the enrichment of genes related to ECM and EMT, as reported in breast and lung cancers (9,27). Although some studies recognized ECM-remodeling CAFs as subtypes of myofibroblasts (6), our results showed that cluster 3 CAFs had a lack of α SMA expression, which indicated a different phenotype from myofibroblasts.

The cluster 4 CAFs exhibited a potential immunoregulation function, which were enriched in antigen processing and presentation. Recent studies have revealed a novel antigen-presenting CAFs (apCAF) in pancreatic ductal adenocarcinoma (28) and breast cancer (29). It was reported that apCAF could express genes related to the major histocompatibility complex (MHC) class II family. *In vivo* examination further confirmed that apCAF could present antigens to T cells. However, the apCAF expressed low level of costimulatory genes, which are required for T cell proliferation. This indicated that apCAF might participate in immunosuppression in the TME (5). The present study revealed that apCAF might also exist in HNSCC tissue.

In summary, we determined the heterogeneity of CAFs

in HNSCC. The prognostic relevance of each CAFs subcluster was estimated, and the potential biological function was determined. However, this study was conducted based on non-high throughput transcriptome data acquired by Smart-Seq technology, in which captured cell numbers were restricted and interference from different patients was unavoidable. Further high throughput single-cell sequencing analysis and *in vivo* study should be carried out for validation.

Acknowledgments

Funding: The authors gratefully acknowledge support for this research from the Jiangsu Province Science & Technology Department (BE2018618), Nanjing Department of Health (YKK18120), Nanjing Medical Science and technique Development Foundation (QRX17174), Nanjing Clinical Research Center for Oral Diseases (No. 2019060009), Jiangsu Postgraduate Research Innovation Program.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at <https://dx.doi.org/10.21037/atm-21-2767>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/atm-21-2767>). All authors report funding from the Jiangsu Province Science & Technology Department (BE2018618), Nanjing Department of Health (YKK18120), Nanjing Medical Science and technique Development Foundation (QRX17174), Nanjing Clinical Research Center for Oral Diseases (No. 2019060009), Jiangsu Postgraduate Research Innovation Program. Dr. HC is from Nanjing Nuoyuan Medical Devices Co., Ltd. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons

Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* 2013;19:1423-37.
2. Sahai E, Astsaturov I, Cukierman E, et al. A framework for advancing our understanding of cancer-associated fibroblasts. *Nat Rev Cancer* 2020;20:174-86.
3. Custódio M, Biddle A, Tavassoli M. Portrait of a CAF: The story of cancer-associated fibroblasts in head and neck cancer. *Oral Oncol* 2020;110:104972.
4. Routray S, Sunkavali A, Bari KA. Carcinoma-associated fibroblasts, its implication in head and neck squamous cell carcinoma: a mini review. *Oral Dis* 2014;20:246-53.
5. Kalluri R. The biology and function of fibroblasts in cancer. *Nat Rev Cancer* 2016;16:582-98.
6. Kanzaki R, Pietras K. Heterogeneity of cancer-associated fibroblasts: Opportunities for precision medicine. *Cancer Sci* 2020;111:2708-17.
7. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med* 2018;50:1-14.
8. Li H, Courtois ET, Sengupta D, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet* 2017;49:708-18.
9. Bartoschek M, Oskolkov N, Bocci M, et al. Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. *Nat Commun* 2018;9:5150.
10. Sebastian A, Hum NR, Martin KA, et al. Single-Cell Transcriptomic Analysis of Tumor-Derived Fibroblasts and Normal Tissue-Resident Fibroblasts Reveals Fibroblast Heterogeneity in Breast Cancer. *Cancers (Basel)* 2020;12:1307.
11. Zhang M, Yang H, Wan L, et al. Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. *J Hepatol* 2020;73:1118-30.
12. Stuart T, Butler A, Hoffman P, et al. Comprehensive

- Integration of Single-Cell Data. *Cell* 2019;177:1888-902.e21.
13. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
 14. Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417-25.
 15. Chen X, Song E. Turning foes to friends: targeting cancer-associated fibroblasts. *Nat Rev Drug Discov* 2019;18:99-115.
 16. Louault K, Li RR, DeClerck YA. Cancer-Associated Fibroblasts: Understanding Their Heterogeneity. *Cancers (Basel)* 2020;12:3108.
 17. Costea DE, Hills A, Osman AH, et al. Identification of two distinct carcinoma-associated fibroblast subtypes with differential tumor-promoting abilities in oral squamous cell carcinoma. *Cancer Res* 2013;73:3888-901.
 18. Puram SV, Tirosh I, Parkh AS, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 2017;171:1611-24.e24.
 19. Yeung TM, Buskens C, Wang LM, et al. Myofibroblast activation in colorectal cancer lymph node metastases. *Br J Cancer* 2013;108:2106-15.
 20. Helms E, Onate MK, Sherman MH. Fibroblast Heterogeneity in the Pancreatic Tumor Microenvironment. *Cancer Discov* 2020;10:648-56.
 21. Öhlund D, Handly-Santana A, Biffi G, et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J Exp Med* 2017;214:579-96.
 22. Liu T, Zhou L, Li D, et al. Cancer-Associated Fibroblasts Build and Secure the Tumor Microenvironment. *Front Cell Dev Biol* 2019;7:60.
 23. Yuzhalin AE, Lim SY, Kutikhin AG, et al. Dynamic matrisome: ECM remodeling factors licensing cancer progression and metastasis. *Biochim Biophys Acta Rev Cancer* 2018;1870:207-28.
 24. Vellinga TT, den Uil S, Rinkes IH, et al. Collagen-rich stroma in aggressive colon tumors induces mesenchymal gene expression and tumor cell invasion. *Oncogene* 2016;35:5263-71.
 25. Provenzano PP, Inman DR, Eliceiri KW, et al. Collagen density promotes mammary tumor initiation and progression. *BMC Med* 2008;6:11.
 26. Penet MF, Kakkad S, Pathak AP, et al. Structure and Function of a Prostate Cancer Dissemination-Permissive Extracellular Matrix. *Clin Cancer Res* 2017;23:2245-54.
 27. Lambrechts D, Wauters E, Boeckx B, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med* 2018;24:1277-89.
 28. Elyada E, Bolisetty M, Laise P, et al. Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer Discov* 2019;9:1102-23.
 29. Costa A, Kieffer Y, Scholer-Dahirel A, et al. Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* 2018;33:463-79.e10.

Cite this article as: Zhang Q, Wang Y, Xia C, Ding L, Pu Y, Hu X, Cai H, Hu Q. Integrated analysis of single-cell RNA-seq and bulk RNA-seq reveals distinct cancer-associated fibroblasts in head and neck squamous cell carcinoma. *Ann Transl Med* 2021;9(12):1017. doi: 10.21037/atm-21-2767