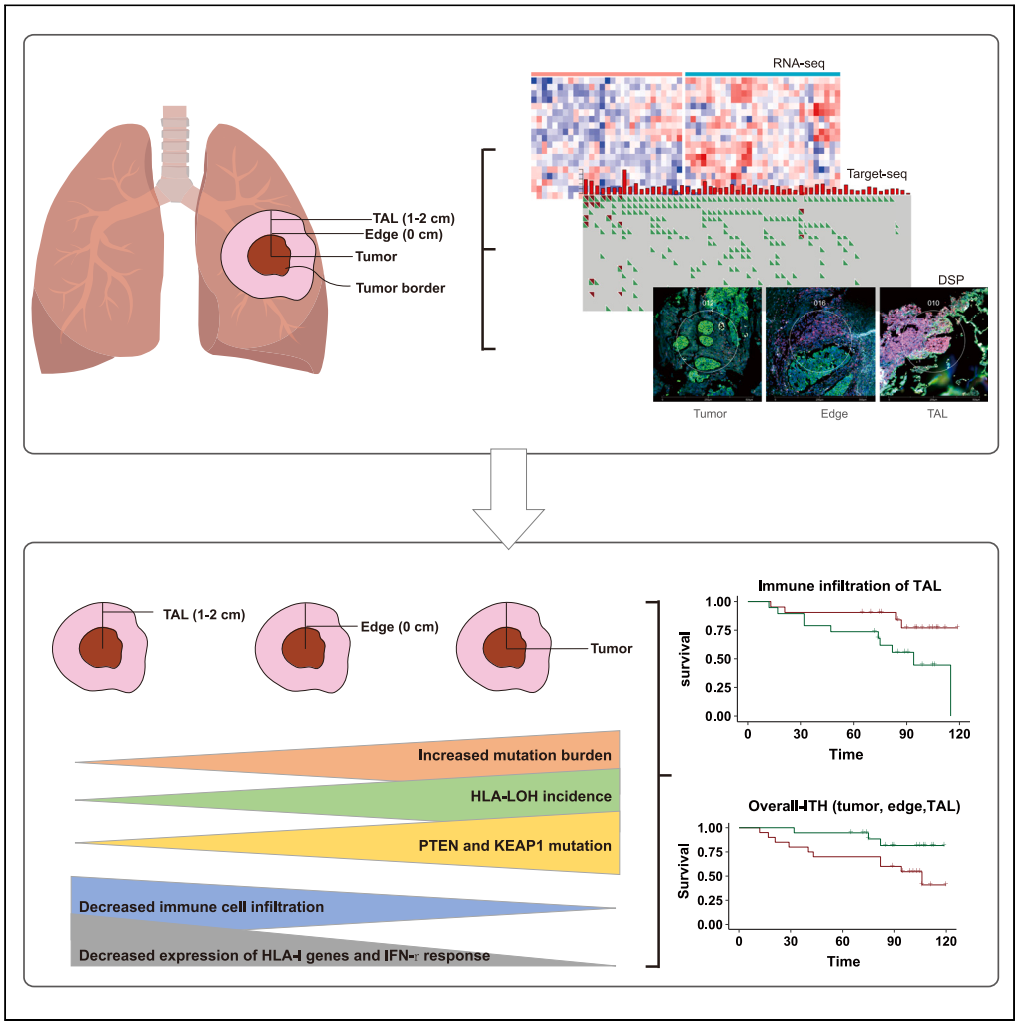
iScience

CellPress
OPEN ACCESS

Article

Immune-active tumor-adjacent tissues are associated with favorable prognosis in stage I lung squamous cell carcinoma



Lisha Ying, Chunliu Zhang, Alexandre Reuben, ..., Xiaodan Pan, Jianjun Zhang, Dan Su

jzhang20@mdanderson.org (J.Z.)
sudan@zjcc.org.cn (D.S.)

Highlights

TALs of stage I LUSC exhibited less immune escape events than tumors

Multiomics analysis indicated more active immune repertoire in TALs

DSP data reveal higher immune heterogeneity score in tumors

Patients with higher immune infiltration in TALs had significantly longer survival

Ying et al., iScience 26, 107732
September 15, 2023 © 2023
https://doi.org/10.1016/j.isci.2023.107732

iScience

CellPress
OPEN ACCESS

Article

Immune-active tumor-adjacent tissues are associated with favorable prognosis in stage I lung squamous cell carcinoma

Lisha Ying,[1,9] Chunliu Zhang,[2,9] Alexandre Reuben,[3] Yiping Tian,[4] Jiaoyue Jin,[4] Canming Wang,[4] Jing Bai,[2] Xinyuan Liu,[4,5] Jianfei Fang,[4] Tingting Feng,[1,4] Chenyang Xu,[4] Rui Zhu,[4] Minran Huang,[4] Yingqi Lyu,[4,6] Tingting Lu,[4,6] Xiaodan Pan,[7] Jianjun Zhang,[3,8,*] and Dan Su[4,10,*]

SUMMARY

The immunogenomic features of tumor-adjacent lungs (TALs) in stage I lung squamous cell carcinoma (LUSC) are not clear. Multiomics analyses of tumor tissues and paired TALs from 59 stage I LUSC patients were performed. Compared to tumors, TALs exhibited a better-preserved immune contexture indicated by upregulation of immune pathways, increased immune infiltration, and higher expression of immune effector molecules. Notably, TALs had no mutations in PTEN and KEAP1, a lower incidence of human leukocyte antigen (HLA) loss and higher expression of HLA class I genes, major histocompatibility complex (MHC) I chaperones, and interferon (IFN)-γ-associated genes. Digital spatial profiling validated the generally higher immune infiltration in TALs and revealed a higher level of immune heterogeneity in LUSC tumors. Importantly, patients with higher immune infiltration in TALs had significantly longer survival, while high immune heterogeneity was associated with inferior patient survival. Our work can be considered in the selection of patients for adjuvant therapy, especially immunotherapy.

INTRODUCTION

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related death worldwide. With the broad implementation of low-dose spiral computed tomography (CT) lung cancer screening, there has been an increase in the diagnosis of early-stage NSCLC. Although early-stage NSCLC can be potentially curable, even for the earliest stage I NSCLC, ~30% of patients still recur.[1] Understanding the mechanisms underlying early carcinogenesis and postsurgical recurrence remains an unmet need to improve lung cancer patient survival.

Accumulating evidence has indicated that molecular and immune features sculpt the evolution of lung cancer. Our previous study showed progressive genomic and epigenetic aberrations from preneoplasia atypical adenomatous hyperplasia (AAH) to invasive lung adenocarcinoma (LUAD), which may collectively impair host immune responses and facilitate immune evasion.[2–4] Compared to LUAD, our understanding of early-stage lung squamous cell carcinoma (LUSC), the second most common subtype of NSCLC,[5–8] remains limited.[9] Furthermore, although immune signatures have been reported to be able to stratify stage non-squamous NSCLCs into high- vs. low-risk groups for recurrence,[10,11] the prognostic signature of LUSC is less defined.

Numerous studies have failed to demonstrate the association of tumor molecular features with postsurgical recurrence of various early-stage cancers.[12–14] One feasible explanation is that molecular profiling of resected tumors cannot completely capture the complex tumor-host interactions that eventually determine postsurgical recurrence. We herein hypothesized that the immune and molecular features of tumor-adjacent lung (TAL) tissues may provide important information on cancer biology and anti-tumor immunity that will impact the prognosis of early-stage LUSC. In this study, we performed deep panel sequencing, RNA sequencing, and digital spatial profiling (DSP) on tumors, tumor edges, and paired TALs from a pure cohort of 59 stage I LUSC patients who were all male smokers to identify molecular and immune features associated with prognosis of stage I LUSC patients.

[1]Zhejiang Cancer Institute, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang 310022, China
[2]Geneplus-Beijing Institute, Beijing, China
[3]Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[4]Department of Pathology, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang 310022, China
[5]The Second Clinical Medical College, Zhejiang Chinese Medicine University, Hangzhou, Zhejiang 310053, China
[6]Department of Oncology, The First Clinical Medical College of Wenzhou Medical University, Wenzhou, Zhejiang 325015, China
[7]Human Tissue Bank, Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang 310022, China
[8]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[9]These authors contributed equally
[10]Lead contact
*Correspondence: jzhang20@mdanderson.org (J.Z.), sudan@zjcc.org.cn (D.S.)
https://doi.org/10.1016/j.isci.2023.107732

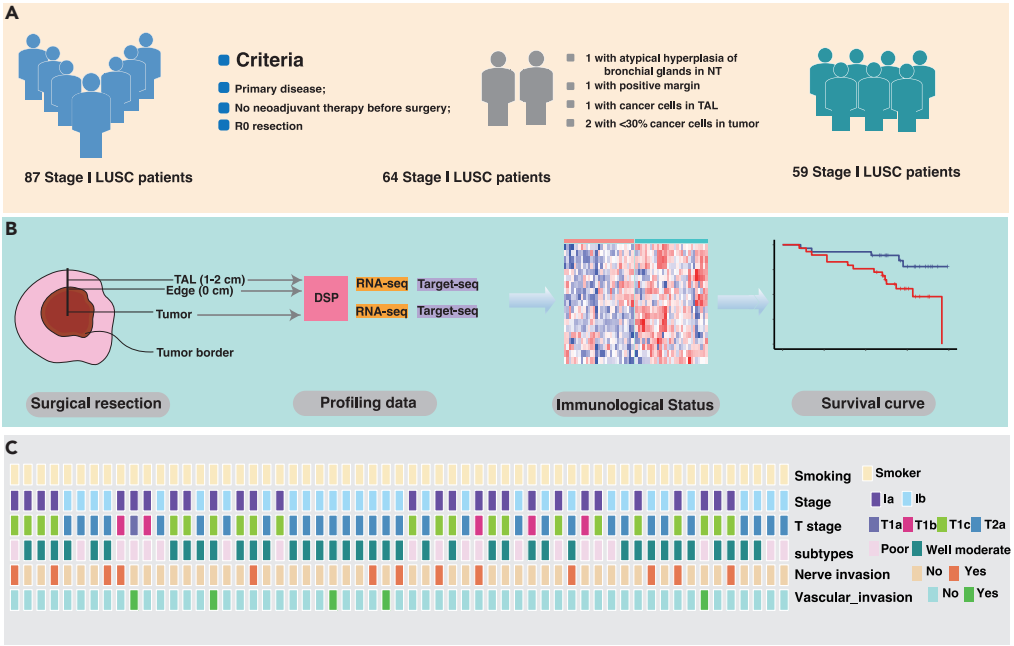


Figure 1. Study design

(A) Schematic of sampling strategy and experimental workflow. NT is for normal tissue.

(B) The tumor tissues, tumor edges (0 cm), and TALs (1–2 cm from tumor border) were resected and subjected to high-throughput sequencing and GeoMx DSP technology. Non-synonymous mutations, transcriptomic profiling, and DSP protein expression data were analyzed to characterize the immune microenvironment and the association between immune microenvironment and survival of stage I LUSC.

(C) Clinical information of every patient enrolled in this study.

RESULTS

Patient characteristics

We retrospectively enrolled 87 stage I LUSC tumors with matched tumor edge and TAL (1–2 cm from the margin). 28 patients were excluded because they received neoadjuvant therapy (n = 21), had R1 resection (n = 2), had hyperplasia in "normal tissues" (n = 1), had positive margin (n = 1), or had cancer cells in TALs (n = 1), or tumor purity was less than 30% (n = 2). This led to a total of 59 tumors that were subjected to multiomics analyses (Figures 1A–B and S1, see details in STAR methods). Clinicopathologic characteristics are listed in Figure 1C, Tables 1, and S2. All patients underwent lobectomy. With a median follow up of 99 months, 29 patients had recurred. At the time of data lockup, 26 patients (40.7%) had died. The median overall survival (OS) was 115 months.

Somatic mutations are present in TALs

To understand the immunogenomic features, 57 tumor tissues and 54 TALs (within 2 cm from the malignant margin) with sufficient tissue were subjected to 1021-gene panel sequencing (Figure S1). A total of 720 non-silent mutations (stop gain, stop loss, frameshift, coding sequence-insertion, coding sequence-deletion, missense, nonsense, etc.) were detected in tumor tissues (Table S3). The most commonly mutated genes included *TP53* (94.7%), *LRP1B* (33.3%), *CDKN2A* (33.3%), and *FAT1* (31.6%) (Figure 2A), consistent with the CHOICE study on Chinese LUSC patients (Figure S2).[15]

A total of 57 non-silent mutations were detected in 15 of 54 TALs including 36 mutations shared in paired tumors and the remaining 21 mutations exclusively detected in the TALs (Figure S3). The dominant mutant genes in TALs were *TP53* (9.3%), *CARD11* (7.4%), *FAT1* (5.6%), *MLL2* (3.7%), and *CDKN2A* (3.7%), among which *TP53* (5 pts), *FAT1* (2 pts), *MLL2* (2 pts), and *CDKN2A* (2 pts) were shared with paired tumors. As expected, the mutation burden in TALs was lower than that in tumor ($p < 2.22e-16$, Wilcoxon test) (Figure 2B). Moreover, mutation burden increased parallel to T stage (Figure 2C). No significant association was found between clinical features and cancer gene mutations identified either in tumors (Figure S4A) or TALs (Figure S4B).

Higher immune infiltration in TALs

As 6 tumor tissues and 17 TAL tissues had insufficient remaining tissue, 53 tumors and 42 TALs were subjected to RNA sequencing to further understand the molecular changes in tumors and TALs. Unsupervised clustering demonstrated that tumor tissues and TALs were clearly separated (Figure S5A). 7,792 differentially expressed genes were identified, of which 3,438 genes were up-regulated and 4,354 genes were

Table 1. Clinicopathological characteristics of stage I LUSC patients

Variable	Cases, n (%)
Age	
≤60 years	20 (33.9%)
>60 years	39 (66.1%)
Sex	
male	59 (100%)
Histology/Grade	
Well/moderate	39 (66.1%)
Poor	20 (33.9%)
Smoking	
Smoker	59 (100%)
Vascular invasion	
Yes	5 (8.0%)
No	54 (92.0%)
Nerve invasion	
Yes	13 (17.2%)
No	46 (82.8%)
T stage	
T1a	1 (1.69%)
T1b	5 (8.47%)
T1c	22 (37.3%)
T2a	31 (52.5%)
Stage	
Ia	28 (47.5%)
Ib	31 (52.5%)

The AJCC (American Joint Committee on Cancer) 8th edition lung cancer staging system was used to classify tumor stage.

down-regulated in tumor tissues (|log2(fold change)|≥1, false discovery rate [FDR]≤0.05, Table S4). Genes up-regulated in tumors were linked to oncogenic pathways (mismatch repair, DNA replication, cell cycle, etc.) and metabolic pathways (alanine, aspartate and glutamate metabolism, biosynthesis of amino acids, metabolism of xenobiotics by cytochrome P450, and glycolysis that is instrumental for angiogenesis, etc.), while those down-regulated in tumors were enriched in immune-associated pathways (antigen processing and presentation, tumor necrosis factor [TNF] signaling pathway, natural killer [NK] cell-mediated cytotoxicity, chemokine signaling pathway, etc.) (Figure 3A) suggesting a more active immune response in TALs than tumors.

We further applied single sample gene set enrichment analysis (ssGSEA)[16] on 28 immune cell signatures to assess the immune cell composition in tumors and TALs (Tables S5 and S6). Overall, both adaptive and innate immune cell infiltration were higher in TALs than tumors except memory B cells, active CD4 cells, CD56 dim NK cells, and CD56 bright NK cells (Figure 3B). Next, we assessed cytolytic activity by evaluating the expression of granzyme A (GZMA) and perforin1 (PRF1)[17] and interferon (IFN)-γ score defined by hallmark gene set from MsigDB.[18] All factors were higher in TALs (Figure 3C), which was confirmed by ESTIMATE analysis[19] (Figure S9A). Taken together, these results suggested a better-preserved immune contexture in TALs compared to tumors.

Molecular basis for immunosuppression in tumors

Next, we attempted to understand the molecular mechanisms underlying immunosuppression in tumors (Figure 4A). Loss of human leukocyte antigen (HLA) heterozygosity (HLALOH), an important immune escape mechanism commonly reported in lung cancers,[20] was observed in 60% of tumor tissues compared to only 9% in TALs (p = 3.696e-08, Fisher's exact test, Figures 4B and S6A). Furthermore, expression of HLA class I genes (HLA-A, HLA-B, HLA-C) was higher in TALs (Figure 4C), as was expression of other major histocompatibility complex (MHC) I chaperones and IFN-γ-associated molecules such as IRF1, IFNGR1, IFNGR2, and PSMB10 (Figure S6B). Overall, these findings indicate a better-preserved antigen presentation machinery and IFN-γ response in TALs. Furthermore, mutations in *KEAP1* and *PTEN* which have been previously associated with immunosuppression were detected in 12.3% (7/57) and 10.5% (6/57) of tumors, respectively, but absent from TALs (Figure 4D). Interestingly, alterations in *PTEN* and HLALOH appeared mutually exclusive, although statistical significance was not

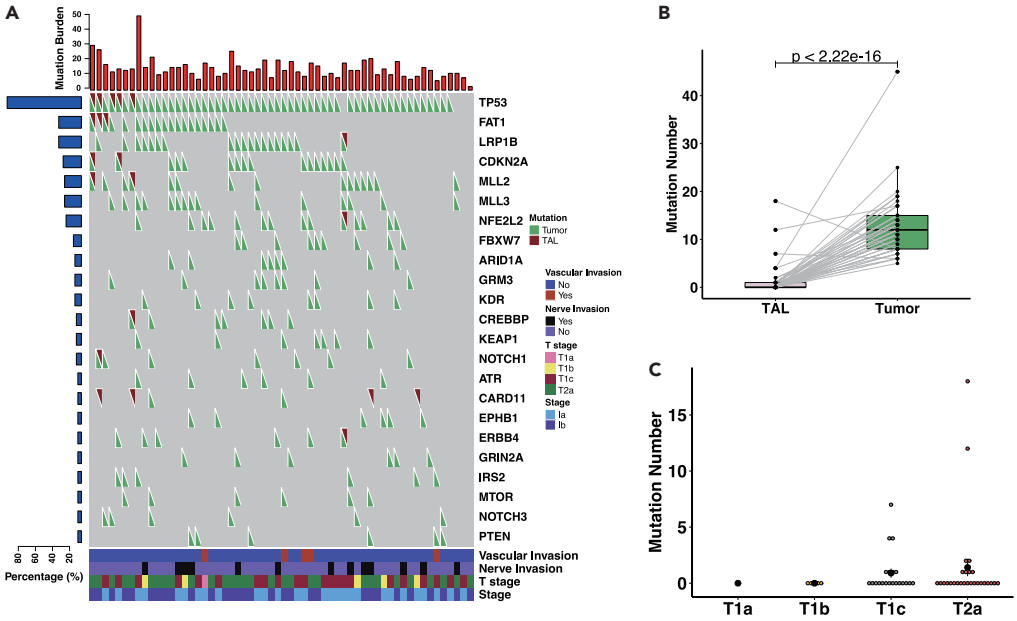


Figure 2. Mutation landscape of tumors and TALs

(A) Mutation landscape in tumors and TALs, respectively (red triangle for TAL, green triangle for tumor). Mutation number of every tumor was on the top of the heatmap. The ratio of every cancer-associated gene was plotted on the left of the heatmap. The clinical characteristics of every patient were plotted on the bottom of the heatmap. And the annotation bar was listed on the right of the heatmap.

(B) The comparison of mutation number between TALs and tumors.

(C) Mutation numbers progressively increased from T1a to T2a in TALs.

reached (p = 0.073, OR = 8.13, Figure 4E). These data reinforce that although immunosuppression is universal in all LUSC tumors, the mechanisms underlying immune evasion may differ in line with the concept of convergent evolution.[21–23]

Multiplexed proteomic profiling demonstrated suppressed and heterogeneous immune repertoire in tumor tissues compared to TALs

To orthogonally validate these intriguing findings in the context of the spatial architecture, we performed DSP, which enables multiplex proteomic profiling of 18 proteins on 42 patients with enough remaining tissues available (Figure 5A). In total, 248 regions of interest (ROIs) within tumors, tumor edges, and TALs were analyzed to infer the spatial organization of the immune microenvironment (Figures 1B and S1). Overall, there was a progressive decrease in immune infiltration as determined by pan-leukocyte marker CD45 from TALs, to tumor edges into the tumor (Figure 5B). In fact, expression of all assessed proteins followed the same trend with a progressive decrease from TALs to tumor edges and tumors (Figures 5C–5E). These results were consistent with the findings from RNA sequencing data demonstrating immune exclusion in stage I LUSC tumors.

Multiple ROIs from geographically separated regions within the same tumor by DSP offered the opportunity to assess immune intratumor heterogeneity (ITH), which may impact overall anti-tumor immunity.[24] As expected, ITH was observed in all immune markers and all tissue types (tumors, tumor edges, and TALs) (Figure S7A). As shown by two representative cases (Figure S7B), each patient demonstrated distinct ITH patterns in tumors, tumor edge, and TALs. Of note, tumors demonstrated higher heterogeneity scores than TALs and tumor edges (Figure 5F), which may further impair anti-tumor immunity in tumor tissues.

Patients with increased infiltration in TALs exhibited improved prognosis

To understand the potential impact of immune features on prognosis, transcriptomic data from TALs and tumor tissues were subjected to unsupervised clustering using the activity score of 28 immune cells (Figures 6A and 6B). Both anti-tumor immune cells (activated CD4, activated CD8, central memory CD4, central memory CD8, effect memory CD4, effect memory CD8, Th1, Th17, activated dentritic cell (DC), CD56briNK, NK, NKT) and pro-tumor immune cells (Treg, Th2, CD56dimNK, immature DC, Myeloid-derived suppressor cells (MDSC), Neutrophil and pDC)[25] were higher in the immune-high group. Pearson correlation analysis showed that the abundance of these two categories of cells was positively associated within the local environment both in tumor ($r = 0.93$, $p < 0.001$, Pearson correlation) and TALs ($r = 0.73$, $p < 0.001$, Pearson correlation, Figures S8A and S8B) respectively reflecting the presence of a feedback mechanism balancing recruitment or differentiation of activating and suppressing immune cells during anti-tumor immune response. Importantly, higher immune infiltration in

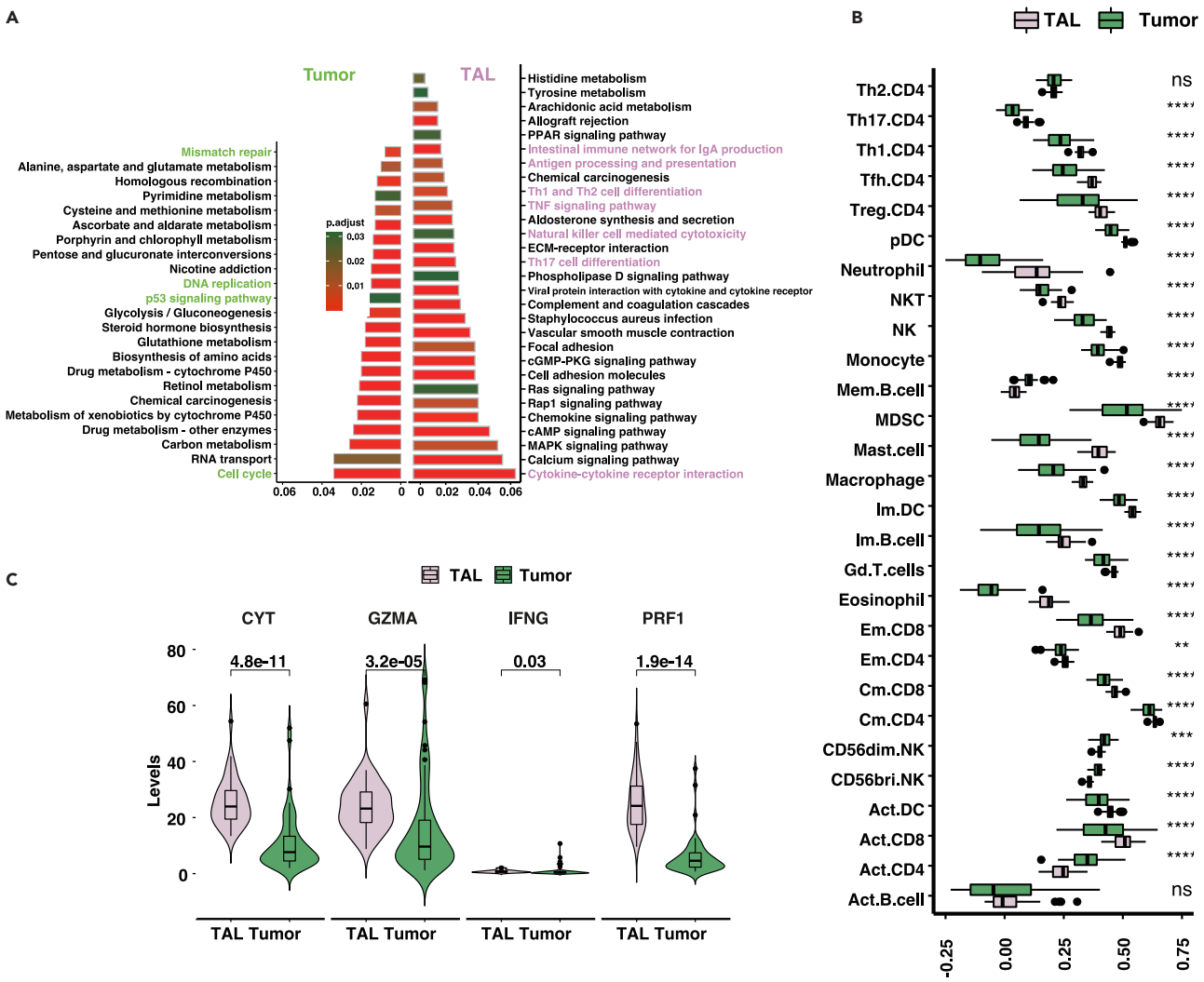


Figure 3. Immune infiltration in tumors and TALs

(A) KEGG pathways were enriched by different expression genes, left for up-regulated genes in tumors and right for up-regulated genes in TALs. The light pink was immune-related pathways in TALs. The light green was oncogenic pathways in tumors.

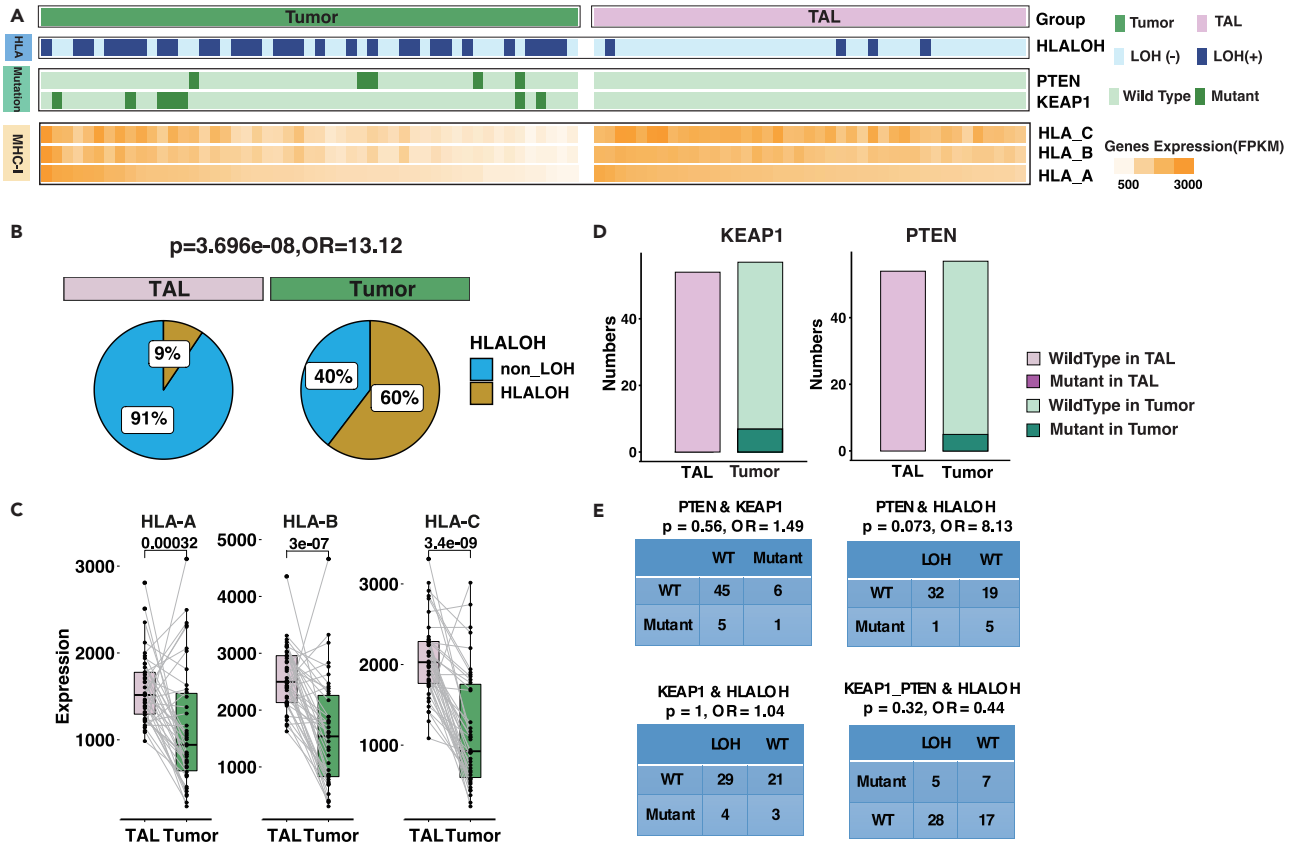
(B) The comparison of activity scores in 28 immune cells between tumors and TALs (* for $p < 0.05$, ** for $p < 0.01$, paired Wilcoxon test).

(C) The comparison of CYT, GZMA, IFN-γ, and PRF1 between tumors and TAL.

TALs but not in tumors was associated with significantly longer OS (OS: $p = 0.036$ for TALs and $p = 0.17$ for tumor; Disease-free survival (DFS): $p = 0.12$ for TALs and $p = 0.79$ for tumor, log rank test, Figures 6C and 6D), in line with results from ESTIMATE (Figures S9B and S9C). Consistently, higher expression of the pan-T cell marker CD3 by DSP analysis in TAL, but not in tumor or tumor edge, was associated with superior survival (Figure 6E). Furthermore, higher immune ITH was associated with significantly inferior OS and a trend of shorter DFS (Figure 6F). This remains significant after multivariate analysis adjusted for age, stage, pleural invasion, vascular invasion, nerve invasion, and tumor differentiation (Table S7).

DISCUSSION

Multiple studies have demonstrated worse prognoses in patients with early-stage LUSC compared to those with LUAD of the same stages.[26–29] Unlike LUAD, very few LUSC tumors carry genomic alterations that are targetable. On the other hand, the majority of LUSC patients are smokers, which has been linked to favorable response to immune checkpoint blockade (ICB) therapy.[30] Accordingly, ICB has brought hope to improve the outcome of LUSC patients. Indeed, immune ICB has demonstrated promising efficacy for early-stage LUSC in both neoadjuvant and adjuvant settings.[31,32] However, not all patients achieve long-term clinical benefit from ICB and currently ICB is



	WT	Mutant
WT	45	6
Mutant	5	1

	LOH	WT
WT	32	19
Mutant	1	5

	LOH	WT
WT	29	21
Mutant	4	3

	LOH	WT
Mutant	5	7
WT	28	17

Figure 4. Immunosuppression factors in tumors and TALs

(A) Landscape of immunosuppression factors in tumors and TALs (Patients performed by target sequencing and RNA sequencing were shown in heatmap).

(B) The occurrence of HLALOH in tumors and TALs ($p < 0.05$, Fisher's exact test).

(C) By comparing the expression level of HLA class I genes (HLA-A, HLA-B, HLA-C) between tumors and TALs, higher expression level was observed in TALs. Expression level was calculated by FPKM.

(D) Mutation ratio of *KEAP1* and *PTEN* in tumors and TALs, respectively.

(E) The association between HLALOH and alterations of *KEAP1* and *PTEN*.

only indicated for LUSC of stages IB–IIIA. As such, understanding the mechanisms underlying postsurgical recurrence and establishing biomarkers to identify high-risk early-stage LUSC patients who may benefit from neoadjuvant or adjuvant ICB-based therapy are warranted.

LUSC tumors have been scrutinized to identify immune and genomic features associated with risk of postsurgical recurrence. Unfortunately, the results were inconsistent and many studies showed no clear correlation[13,14,33,34] reflecting the complexity of host-tumor interactions that may determine postsurgical recurrence. On the other hand, TALs, which may provide critical information on host-tumor interactions, have not been systemically investigated. In the current study, we characterized the molecular and immune profiles of tumors and TALs from 59 patients with stage I LUSC. Interestingly, 57 somatic mutations were identified in TALs, in line with findings from urothelial carcinoma[35] suggesting that epithelial cells in TALs have already undergone cancerous transformation at the molecular level, despite being pathologically non-malignant.

Interestingly, we observed a progressive change of immune repertoire from TALs to tumor edges and LUSC tumors with an overall decrease in anti-tumor responses and an increase in infiltration of suppressive or dysfunctional immune cells in LUSC tumors. Cancer gene mutations associated with immunosuppression such as mutations in *PTEN* and *KEAP1*,[36,37] HLA loss, and downregulation of immune genes (GZMA, IFN-γ, PRF1, etc.) may be the molecular features underlying immune evasion in LUSC tumors. Importantly, higher immune infiltration in TALs but not in tumors was associated with superior survival. These results suggest that the immune microenvironment in TALs may better reflect the local anti-tumor immune response than tumor tissues and thus should be taken into account when assessing the risk of postoperative recurrence of early-stage LUSC. Furthermore, by utilizing DSP analysis, we investigated the spatial immune infiltration in tumors, tumor edges, and TALs. Although immune heterogeneity was present across all tissue types as expected, LUSC tumors exhibited the highest level of immune heterogeneity, which may further facilitate immune evasion of LUSC tumor cells. Importantly, higher overall immune heterogeneity was associated with inferior patient survival, consistent with prior studies.[38]

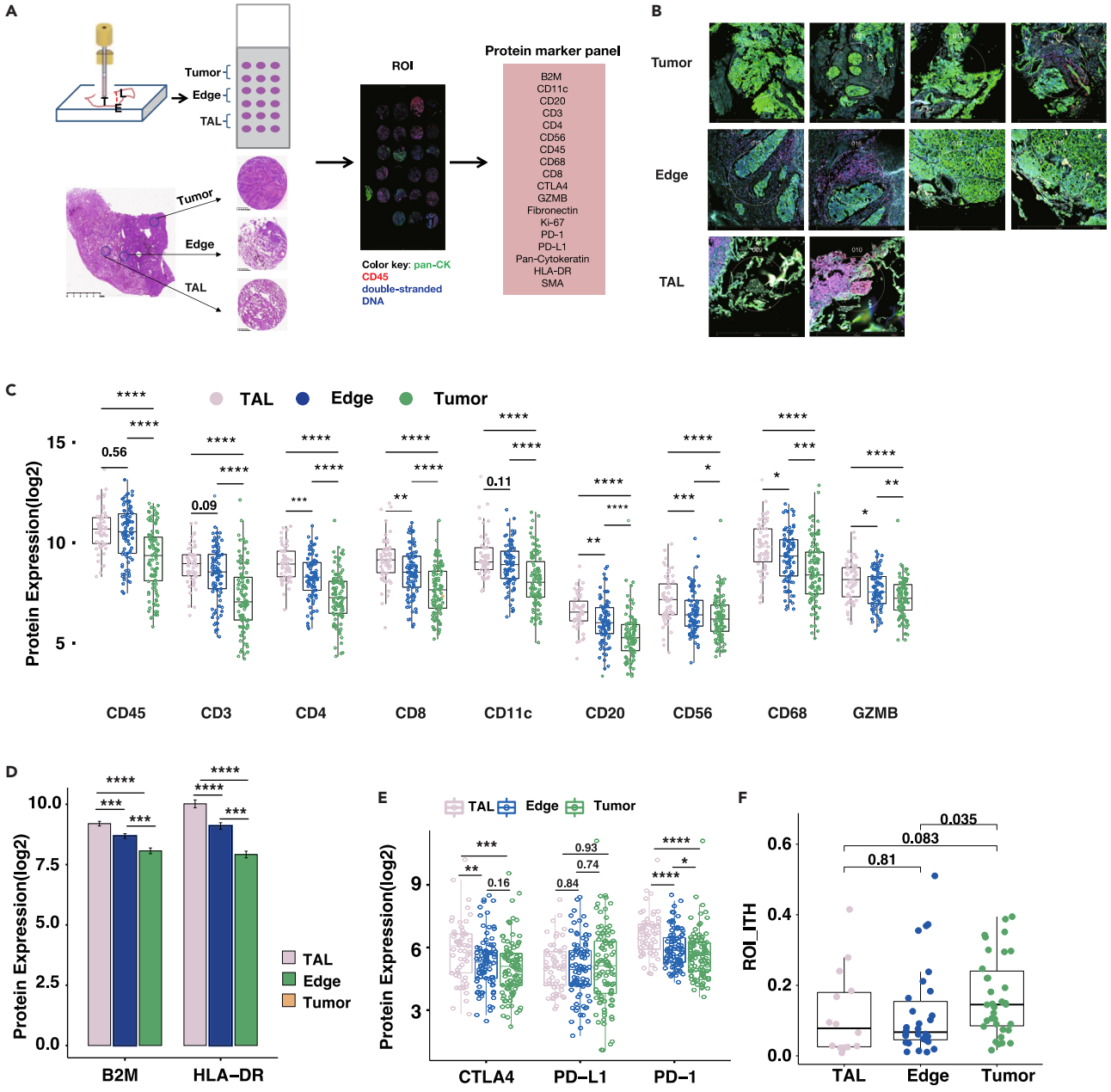


Figure 5. Better immune cell infiltration in TALs was validated by DSP data

(A) Multiple ROIs per tissue were selected based on pan-CK and CD45 and subjected to spatial proteomic profiling of 21 tumor and immune markers.

(B) Representative images of ROIs from P02 and quantification of CD45 protein levels in tumor, tumor edge, and TAL regions.

(C) Comparison of expression level of DSP immune markers across TALs, tumor edges, and tumors.

(D) Comparison of expression level of B2M and HLA-DR across TALs, tumor edges, and tumors.

(E) Boxplot showing difference of immune checkpoint biomarker of CTLA4, PD-L1, and PD-1 in tumors, tumor edges, and TALs.

(F) The comparison of heterogeneity score across TALs, tumor edge, and tumors. Immune heterogeneity score was assessed by Max (paired-ROIs cosin similarity) - Min(paired-ROIs cosin similarity). (p values are based on a two-sided Wilcoxon test).

Our results suggest that TALs may provide critical information for host anti-tumor immune response and prognosis. Therefore, profiling of both tumor tissues and TALs should be considered to accurately identify early-stage LUSC patients at high risk for postsurgical recurrence. Since higher immune infiltration in TALs but not in tumors was associated with superior survival, we anticipate incorporating TALs into clinical decision-making may provide additional insights in clinical practice. From a translational perspective, TAL and tumor

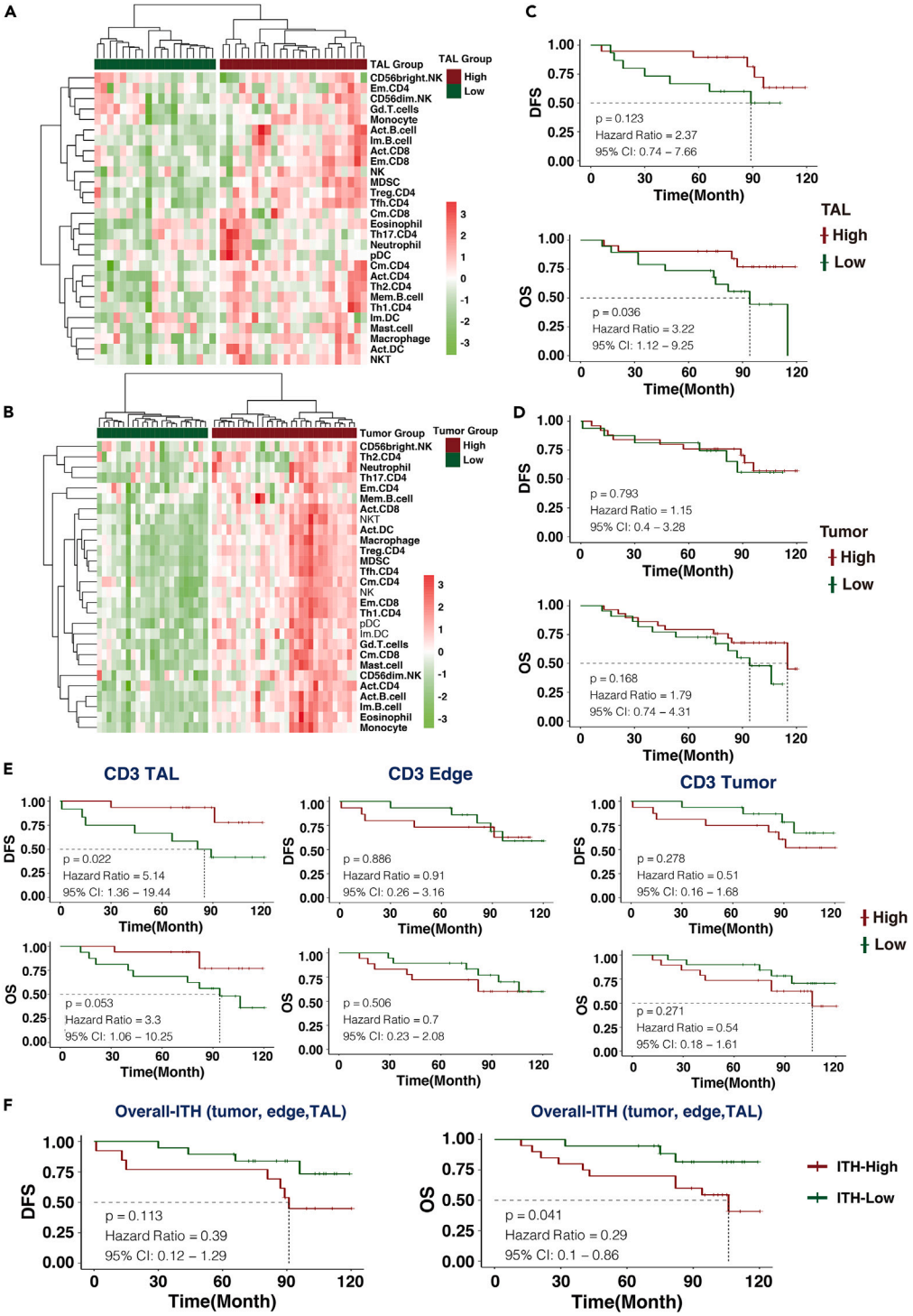


Figure 6. The association of Immune infiltration groups and survival

(A and B) Two immune subgroups were evaluated by unsupervised cluster of 28 immune cells in TALs (A) and tumors (B), respectively.

(C and D) The survival curves of immune group in TALs, and tumors (red for high immune infiltration group, green for low immune infiltration group).

(E) Kaplan-Meier graphs showing DSP CD3 protein in TALs was associated with survival, while irrelevant to outcomes in tumor edge and tumors.

(F) The association between overall immune ITH and survival. The overall immune ITH level was evaluated by average ITH score of tumor, tumor edge, and TAL in one patient.

infiltration can be assessed by immunohistochemistry (IHC) or gene expression profiling, which may help identify high-risk patients who may benefit from adjuvant immunotherapy.

Conclusions

Our study provides an insight into the immune microenvironment of TALs in stage I LUSC patients and its association with patient survival, which may contribute to identify high-risk patients who may benefit from neoadjuvant or adjuvant ICB therapy.

Limitations of the study

As a retrospective study, our work has several important limitations. First, in order to minimize confounding factors that may impact molecular/immune features as well as survival, we focused on a very uniform patient population: stage I LUSC patients with no neoadjuvant or adjuvant treatment, who were all male and smokers. Therefore, whether these intriguing findings are broadly applicable to other patient populations remains to be determined. Second, many stage I LUSC specimens were inadequate for multiomics profiling due to small sample size that limited our ability to reach statistical significance in many of our analyses. Additional experiments to validate the molecular changes observed in tumors and TALs were warranted. Third, retrospectively collected frozen specimens limited our ability to perform high-resolution assays such as single-cell profiling to pinpoint the specific subset of immune cells relevant to postsurgical recurrence. Nevertheless, as a proof-for-principle study, our work reveals the molecular landscape of TALs and suggested that TALs can provide critical information regarding local anti-tumor immune response that may be important for postsurgical recurrence. In light of recent approval of ICB in the neoadjuvant setting,[39] molecular and immune profiling of TALs in addition to tumors after neoadjuvant ICB treatment, ideally with high-resolution technologies, such as single-cell sequencing, may provide valuable insights into mechanisms of resistance to ICB and postsurgical recurrence of LUSC.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Study design and subjects
  - Ethic statement
- METHOD DETAILS
  - DNA and RNA extraction
  - Targeted capture and next-generation sequencing analysis
  - Mutation analysis
  - RNA sequencing data analysis
  - HLA typing and loss of heterozygosity
  - DSP data generation and analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107732.

ACKNOWLEDGMENTS

This study was supported by grants from the Major Science and Technology Project of Zhejiang Province of China (No. 2020C03023), the National Natural Science Foundation of China (No. 82003188), the National Key Research R&D Program of China (No.2021YFA0910100), the Zhejiang Leading Innovation and Entrepreneurship Team (No.2022R01006), Zhejiang Provincial Natural Science Foundation of China under Grant (No.LY22H160025), Zhejiang high-level innovative talent program, and the 1022 program of Zhejiang Cancer Hospital.

AUTHORS CONTRIBUTIONS

DS and JZ conceived and designed the study. LY, JF, TF, CX, RZ, MH, YL, XL, TL, and XP collected samples as well as clinical information. YT, JJ, and CW reviewed the pathological samples. LY and CZ performed the experiments. LY, CZ, JB, JF, and JZ analyzed the data. LY, CZ, AR, JZ, and DS wrote the manuscript. LY, CZ, JZ, and DS provided intellectual discussions and ideas regarding the content of manuscript. JB and DS supervised the study. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

Dr. Reuben serves on the scientific advisory board and has received honoraria from Adaptive Biotechnologies. Dr. Zhang reports grants from Merck, Novartis, Johnson and Johnson, personal fees from BMS, AZ, Novartis, Johnson and Johnson, GenePlus, Innovent, outside the submitted work.

Received: March 5, 2023
Revised: July 7, 2023
Accepted: August 21, 2023
Published: August 25, 2023

REFERENCES

1. Brierley, J.D., Gospodarowicz, M.K., and Wittekind, C. (2017). TNM Classification of Malignant Tumours (John Wiley & Sons).
2. Hu, X., Fujimoto, J., Ying, L., Fukuoka, J., Ashizawa, K., Sun, W., Reuben, A., Chow, C.W., McGranahan, N., Chen, R., et al. (2019). Multi-region exome sequencing reveals genomic evolution from preneoplasia to lung adenocarcinoma. Nat. Commun. *10*, 2978. https://doi.org/10.1038/s41467-019-10877-8.
3. Hu, X., Estecio, M.R., Chen, R., Reuben, A., Wang, L., Fujimoto, J., Carrot-Zhang, J., McGranahan, N., Ying, L., Fukuoka, J., et al. (2021). Evolution of DNA methylome from precancerous lesions to invasive lung adenocarcinomas. Nat. Commun. *12*, 687. https://doi.org/10.1038/s41467-021-20907-z.
4. Dejima, H., Hu, X., Chen, R., Zhang, J., Fujimoto, J., Parra, E.R., Haymaker, C., Hubert, S.M., Duose, D., Solis, L.M., et al. (2021). Immune evolution from preneoplasia to invasive lung adenocarcinomas and underlying molecular features. Nat. Commun. *12*, 2722. https://doi.org/10.1038/s41467-021-22890-x.
5. Niu, Z., Jin, R., Zhang, Y., and Li, H. (2022). Signaling pathways and targeted therapies in lung squamous cell carcinoma: mechanisms and clinical trials. Signal Transduct. Target. Ther. *7*, 353. https://doi.org/10.1038/s41392-022-01200-x.
6. Chen, Z., Fillmore, C.M., Hammerman, P.S., Kim, C.F., and Wong, K.K. (2014). Non-small-cell lung cancers: a heterogeneous set of diseases. Nat. Rev. Cancer *14*, 535–546. https://doi.org/10.1038/nrc3775.
7. Gandara, D.R., Hammerman, P.S., Sos, M.L., Lara, P.N., Jr., and Hirsch, F.R. (2015). Squamous cell lung cancer: from tumor genomics to cancer therapeutics. Clin. Cancer Res. *21*, 2236–2243. https://doi.org/10.1158/1078-0432.CCR-14-3039.
8. Derman, B.A., Mileham, K.F., Bonomi, P.D., Batus, M., and Fidler, M.J. (2015). Treatment of advanced squamous cell carcinoma of the lung: a review. Transl. Lung Cancer Res. *4*, 524–532. https://doi.org/10.3978/j.issn.2218-6751.2015.06.07.
9. Cardarella, S., and Johnson, B.E. (2013). The impact of genomic changes on treatment of lung cancer. Am. J. Respir. Crit. Care Med. *188*, 770–775. https://doi.org/10.1164/rccm.201305-0843PP.
10. Li, B., Cui, Y., Diehn, M., and Li, R. (2017). Development and Validation of an Individualized Immune Prognostic Signature in Early-Stage Nonsquamous Non-Small Cell Lung Cancer. JAMA Oncol. *3*, 1529–1537. https://doi.org/10.1001/jamaoncol.2017.1609.
11. Wu, P., Zheng, Y., Wang, Y., Wang, Y., and Liang, N. (2020). Development and validation of a robust immune-related prognostic signature in early-stage lung adenocarcinoma. J. Transl. Med. *18*, 380. https://doi.org/10.1186/s12967-020-02545-z.
12. Yu, H., Chen, Z., Ballman, K.V., Watson, M.A., Govindan, R., Lanc, I., Beer, D.G., Bueno, R., Chirieac, L.R., Chui, M.H., et al. (2019). Correlation of PD-L1 Expression with Tumor Mutation Burden and Gene Signatures for Prognosis in Early-Stage Squamous Cell Lung Carcinoma. J. Thorac. Oncol. *14*, 25–36. https://doi.org/10.1016/j.jtho.2018.09.006.
13. Stewart, P.A., Welsh, E.A., Slebos, R.J.C., Fang, B., Izumi, V., Chambers, M., Zhang, G., Cen, L., Pettersson, F., Zhang, Y., et al. (2019). Proteogenomic landscape of squamous cell lung cancer. Nat. Commun. *10*, 3578. https://doi.org/10.1038/s41467-019-11452-x.
14. Choi, M., Kadara, H., Zhang, J., Parra, E.R., Rodriguez-Canales, J., Gaffney, S.G., Zhao, Z., Behrens, C., Fujimoto, J., Chow, C., et al. (2017). Mutation profiles in early-stage lung squamous cell carcinoma with clinical follow-up and correlation with markers of immune function. Ann. Oncol. *28*, 83–89. https://doi.org/10.1093/annonc/mdw437.
15. Zhang, X.C., Wang, J., Shao, G.G., Wang, Q., Qu, X., Wang, B., Moy, C., Fan, Y., Albertyn, Z., Huang, X., et al. (2019). Comprehensive genomic and immunological characterization of Chinese non-small cell lung cancer patients. Nat. Commun. *10*, 1772. https://doi.org/10.1038/s41467-019-09762-1.
16. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinf. *14*, 7. https://doi.org/10.1186/1471-2105-14-7.
17. Rooney, M.S., Shukla, S.A., Wu, C.J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell *160*, 48–61. https://doi.org/10.1016/j.cell.2014.12.033.
18. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425. https://doi.org/10.1016/j.cels.2015.12.004.
19. Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat. Commun. *4*, 2612. https://doi.org/10.1038/ncomms3612.
20. McGranahan, N., Rosenthal, R., Hiley, C.T., Rowan, A.J., Watkins, T.B.K., Wilson, G.A., Birkbak, N.J., Veeriah, S., Van Loo, P., Herrero, J., et al. (2017). Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell *171*, 1259–1271.e11. https://doi.org/10.1016/j.cell.2017.10.001.
21. Liu, Y., Zhang, J., Li, L., Yin, G., Zhang, J., Zheng, S., Cheung, H., Wu, N., Lu, N., Mao, X., et al. (2016). Genomic heterogeneity of multiple synchronous lung cancer. Nat. Commun. *7*, 13200. https://doi.org/10.1038/ncomms13200.
22. Chen, M., Chen, R., Jin, Y., Li, J., Hu, X., Zhang, J., Fujimoto, J., Hubert, S.M., Gay, C.M., Zhu, B., et al. (2021). Cold and heterogeneous T cell repertoire is associated with copy number aberrations and loss of immune genes in small-cell lung cancer. Nat. Commun. *12*, 6655. https://doi.org/10.1038/s41467-021-26821-8.
23. Lee, W.C., Reuben, A., Hu, X., McGranahan, N., Chen, R., Jalali, A., Negrao, M.V., Hubert, S.M., Tang, C., Wu, C.C., et al. (2020). Multiomics profiling of primary lung cancers and distant metastases reveals immunosuppression as a common characteristic of tumor cells with metastatic plasticity. Genome Biol. *21*, 271. https://doi.org/10.1186/s13059-020-02175-0.
24. Reuben, A., Gittelman, R., Gao, J., Zhang, J., Yusko, E.C., Wu, C.J., Emerson, R., Zhang, J., Tipton, C., Li, J., et al. (2017). TCR Repertoire Intratumor Heterogeneity in Localized Lung Adenocarcinomas: An Association with Predicted Neoantigen Heterogeneity and Postsurgical Recurrence. Cancer Discov. *7*, 1088–1097. https://doi.org/10.1158/2159-8290.CD-17-0256.
25. Jia, Q., Wu, W., Wang, Y., Alexander, P.B., Sun, C., Gong, Z., Cheng, J.N., Sun, H., Guan, Y., Xia, X., et al. (2018). Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. Nat. Commun. *9*, 5361. https://doi.org/10.1038/s41467-018-07767-w.
26. Tian, S. (2017). Classification and survival prediction for early-stage lung adenocarcinoma and squamous cell carcinoma patients. Oncol. Lett. *14*, 5464–5470. https://doi.org/10.3892/ol.2017.6835.
27. Fukui, T., Taniguchi, T., Kawaguchi, K., Fukumoto, K., Nakamura, S., Sakao, Y., and Yokoi, K. (2015). Comparisons of the clinicopathological features and survival outcomes between lung cancer patients with adenocarcinoma and squamous cell carcinoma. Gen. Thorac. Cardiovasc. Surg. *63*, 507–513. https://doi.org/10.1007/s11748-015-0564-5.
28. Usui, S., Minami, Y., Shiozawa, T., Iyama, S., Satomi, K., Sakashita, S., Sato, Y., and Noguchi, M. (2013). Differences in the prognostic implications of vascular invasion between lung adenocarcinoma and squamous cell carcinoma. Lung Cancer *82*,

407–412. https://doi.org/10.1016/j.lungcan.2013.09.001.

29. Ogawa, H., Uchino, K., Tanaka, Y., Shimizu, N., Okuda, Y., Tane, K., Hokka, D., Tane, S., Tauchi, S., Nishio, W., et al. (2015). Outcomes of segmentectomy for cT1bN0M0 lung adenocarcinoma and squamous cell carcinoma: a possible association with pathological invasion. Eur. J. Cardio. Thorac. Surg. *48*, 77–82. https://doi.org/10.1093/ejcts/ezu429.
30. Gainor, J.F., Rizvi, H., Jimenez Aguilar, E., Skoulidis, F., Yeap, B.Y., Naidoo, J., Khosrowjerdi, S., Mooradian, M., Lydon, C., Illei, P., et al. (2020). Clinical activity of programmed cell death 1 (PD-1) blockade in never, light, and heavy smokers with non-small-cell lung cancer and PD-L1 expression >/=50. Ann. Oncol. *31*, 404–411. https://doi.org/10.1016/j.annonc.2019.11.015.
31. Altorki, N.K., McGraw, T.E., Borczuk, A.C., Saxena, A., Port, J.L., Stiles, B.M., Lee, B.E., Sanfilippo, N.J., Scheff, R.J., Pua, B.B., et al. (2021). Neoadjuvant durvalumab with or without stereotactic body radiotherapy in patients with early-stage non-small-cell lung cancer: a single-centre, randomised phase 2 trial. Lancet Oncol. *22*, 824–835. https://doi.org/10.1016/S1470-2045(21)00149-2.
32. Shu, C.A., Gainor, J.F., Awad, M.M., Chiuzan, C., Grigg, C.M., Pabani, A., Garofano, R.F., Stoopler, M.B., Cheng, S.K., White, A., et al. (2020). Neoadjuvant atezolizumab and chemotherapy in patients with resectable non-small-cell lung cancer: an open-label, multicentre, single-arm, phase 2 trial. Lancet Oncol. *21*, 786–795. https://doi.org/10.1016/S1470-2045(20)30140-6.
33. Satpathy, S., Krug, K., Jean Beltran, P.M., Savage, S.R., Petralia, F., Kumar-Sinha, C., Dou, Y., Reva, B., Kane, M.H., Avanessian, S.C., et al. (2021). A proteogenomic portrait of lung squamous cell carcinoma. Cell *184*, 4348–4371.e40. https://doi.org/10.1016/j.cell.2021.07.016.
34. Jiang, T., Shi, J., Dong, Z., Hou, L., Zhao, C., Li, X., Mao, B., Zhu, W., Guo, X., Zhang, H., et al. (2019). Genomic landscape and its correlations with tumor mutational burden, PD-L1 expression, and immune cells infiltration in Chinese lung squamous cell carcinoma. J. Hematol. Oncol. *12*, 75. https://doi.org/10.1186/s13045-019-0762-1.
35. Li, R., Du, Y., Chen, Z., Xu, D., Lin, T., Jin, S., Wang, G., Liu, Z., Lu, M., Chen, X., et al. (2020). Macroscopic somatic clonal expansion in morphologically normal human urothelium. Science *370*, 82–89. https://doi.org/10.1126/science.aba7300.
36. Liao, W., Overman, M.J., Boutin, A.T., Shang, X., Zhao, D., Dey, P., Li, J., Wang, G., Lan, Z., Li, J., et al. (2019). KRAS-IRF2 axis drives immune suppression and immune therapy resistance in colorectal cancer. Cancer Cell *35*, 559–572.e7. https://doi.org/10.1016/j.ccell.2019.02.008.
37. Marinelli, D., Mazzotta, M., Scalera, S., Terrenato, I., Sperati, F., D'Ambrosio, L., Pallocca, M., Corleone, G., Krasniqi, E., Pizzuti, L., et al. (2020). KEAP1-driven co-mutations in lung adenocarcinoma unresponsive to immunotherapy despite high tumor mutational burden. Ann. Oncol. *31*, 1746–1754. https://doi.org/10.1016/j.annonc.2020.08.2105.
38. Nguyen, P.H.D., Ma, S., Phua, C.Z.J., Kaya, N.A., Lai, H.L.H., Lim, C.J., Lim, J.Q., Wasser, M., Lai, L., Tam, W.L., et al. (2021). Intratumoural immune heterogeneity as a hallmark of tumour evolution and progression in hepatocellular carcinoma. Nat. Commun. *12*, 227. https://doi.org/10.1038/s41467-020-20171-7.
39. Forde, P.M., Spicer, J., Lu, S., Provencio, M., Mitsudomi, T., Awad, M.M., Felip, E., Broderick, S.R., Brahmer, J.R., Swanson, S.J., et al. (2022). Neoadjuvant Nivolumab plus Chemotherapy in Resectable Lung Cancer. N. Engl. J. Med. *386*, 1973–1985. https://doi.org/10.1056/NEJMoa2202170.
40. Chen, K., Bai, J., Reuben, A., Zhao, H., Kang, G., Zhang, C., Qi, Q., Xu, Y., Hubert, S., Chang, L., et al. (2021). Multiomics Analysis Reveals Distinct Immunogenomic Features of Lung Cancer with Ground-Glass Opacity. Am. J. Respir. Crit. Care Med. *204*, 1180–1192. https://doi.org/10.1164/rccm.202101-0119OC.
41. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
42. Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. *31*, 213–219. https://doi.org/10.1038/nbt.2514.
43. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303. https://doi.org/10.1101/gr.107524.110.
44. Li, J., Lupat, R., Amarasinghe, K.C., Thompson, E.R., Doyle, M.A., Ryland, G.L., Tothill, R.W., Halgamuge, S.K., Campbell, I.G., and Gorringe, K.L. (2012). CONTRA: copy number analysis for targeted resequencing. Bioinformatics *28*, 1307–1313. https://doi.org/10.1093/bioinformatics/bts146.
45. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192. https://doi.org/10.1093/bib/bbs017.
46. Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360. https://doi.org/10.1038/nmeth.3317.
47. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. *33*, 290–295. https://doi.org/10.1038/nbt.3122.
48. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169. https://doi.org/10.1093/bioinformatics/btu638.
49. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.
50. Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl, H., and Trajanoski, Z. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep. *18*, 248–262. https://doi.org/10.1016/j.celrep.2016.12.019.
51. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.
52. Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics *30*, 3310–3316. https://doi.org/10.1093/bioinformatics/btu548.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
PanCK(Clone:AE-1/AE-3)	Novus Biologicals	Cat# NBP2-33200
CD45(Clone:2B11+PD7/26)	Novus Biologicals	Cat# NBP2-34528
Syto13	Thermo Fisher Scientific	Cat# S7575
Biological samples		
LUSC patients frozen tumor, tumor edge and tumor-adjacent lungs	Zhejiang Cancer Hospotal	N/A
Deposited data		
targeted-panel sequencing and RNA sequencing data	GSA for Human (https://ngdc.cncb.ac.cn/gsa-human/)	PRJCA010105
Software and algorithms		
Quality Control of Fastq	N/A	Customized Scripts
BWA	v0.7.12-r1039	https://bio-bwa.sourceforge.net/
Picard	v1.98	https://broadinstitute.github.io/picard/
GATK	v3.4-46-gbc02625	https://gatk.broadinstitute.org/hc/en-us
MuTect	v1.1.4	https://gatk.broadinstitute.org/hc/en-us/articles/360037593851-Mutect2
CONTRA	v2.0.8	https://contra-cnv.sourceforge.net/
HISAT2	v2.1.0	http://daehwankimlab.github.io/hisat2/
StringTie	v1.2.3	https://ccb.jhu.edu/software/stringtie/
HTSeq	v0.6.0	https://htseq.readthedocs.io/en/master/
DESeq2	v1.40.2	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
clusterProfiler	V4.8.2	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
ssGSEA	v10.0.7	https://gsea-msigdb.github.io/ssGSEA-gpmodule/v10/index.html
Optitype	v1.3.3	https://github.com/FRED-2/OptiType
LOH HLA	N/A	https://github.com/mskcc/lohhla
R	v3.6.3	https://www.r-project.org/
Python	v3.7.10	https://www.python.org/downloads/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Prof. DanSu (sudan@zjcc.org.cn).

Materials availability

This study did not generate new unique materials.

Data and code availability

- The processed targeted-panel sequencing, RNA sequencing and DSP data in this study have been deposited in the Genome Sequence Archive for Human (GSA-Human; https://ngdc.cncb.ac.cn/gsa-human/), a data repository specialized for storing human genetic data derived from biomedical research. Data are available under accession number GSA-Human: PRJCA010105 upon formal request. However,

to protect patient privacy, interested researchers may need to apply via a data access committee, which will evaluate all requests made by applicants and grant access if deemed reasonable. All other data may be found within the manuscript or supplementary information.

- Code is available from the lead contact upon request.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study design and subjects

87 heavy smokers (all male) with clinical stage I LUSC treated at the Zhejiang Cancer Hospital in China between April 2008 and March 2015, who had frozen tumor, tumor edge and TAL (1–2 cm from the margin) available were queried. 23 patients were excluded because they received neoadjuvant therapy or had R1 resection. All tumors and TALs were subjected to extensive pathology review and 5 patients were further excluded including one patient with atypical hyperplasia of bronchial glands in normal tissues, one patient with positive margin, one patient with cancer cells in TALs and two patients with <30% cancer cells in their tumors (Figure 1A). All tumor samples had a minimum of 30% tumor cells. All patients were followed up until time of death or censored following surgery. Time of follow up was interrupted for this study on January 9, 2020. The median follow-up time was 99 months. This study was approved by the institutional review board of the Zhejiang cancer Hospital. Tumor tissues, tumor edges, TALs and leukocytes were subjected to targeted sequencing, RNA sequencing and DSP (Figures 1B and S1A). Two representative cases of H&E and immunohistochemistry images in tumor tissues and TALs were listed in Figure S1B. Clinical data, including age, sex, developmental stage, was extracted from medical records and listed in Tables 1 and S2.

Ethic statement

This study was approved by The Zhejiang Cancer Hospital (IRB-2020-63). The informed consent form was signed by every participant, allowing analysis of blood samples, tumor tissues, TALs and clinical data. The study was designed and performed according to the principles of the Declaration of Helsinki.

METHOD DETAILS

DNA and RNA extraction

Genomic DNA (gDNA) was extracted from tumors and tumor-adjacent tissues and leukocyte layer using the QIAamp DNA MiniKit and QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany), respectively, according to the manufacturer's instructions. DNA from leukocyte layer was used as germline control. The Qubit fluorometer and dsDNA HS (High Sensitivity) Assay Kit (Invitrogen, Carlsbad, CA, USA) were used to measure DNA concentration. The Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) was utilized to assess the distribution of DNA. RNA was extracted from tissues using Trizol and RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany).

Targeted capture and next-generation sequencing analysis

Sequencing libraries of gDNA were constructed with the KAPA DNA Library Preparation Kit (Kapa Biosystems, Wilmington, MA, USA), followed the manufacturer's protocol. Libraries were hybridized to custom-designed biotinylated oligonucleotide probes (Integrated DNA Technologies, Iowa, IA, USA). The gDNA from tumors and tumor-adjacent tissues and leukocyte were subjected to next-generation sequencing of 1,021 frequently mutated cancer genes and exons 2–3 of HLA genes (see Table S1 in the online supplement).[40] Next generation sequencing (NGS) was performed using the HiSeq 2000 Sequencing System (Illumina, San Diego, CA) with 2 × 101-bp paired-end reads. DNA from leukocyte was used as germline control.

The quality control process was applied in FASTQ data from tumors and TALs and control sample by removing the terminal adaptor sequences and low-quality reads from raw data. BWA[41] (version 0.7.12-r1039) was employed to align the clean reads to the reference human genome (hg19). Picard (version 1.98) was used to mark PCR duplicates. Realignment and recalibration were performed for these redundant reads using GATK (version 3.4–46-gbc02625).

Mutation analysis

Single nucleotide variants (SNV) were called by MuTect[42] (version 1.1.4) and NChot, a software developed in-house to review hotspot variants. Small insertions and deletions (Indels) were called by GATK.[43] Somatic copy-number alterations were identified with CONTRA[44] (v2.0.8). Significant copy number variations were expressed as the ratio of adjusted depth between tumor DNA and control germline DNA. The final candidate variants were all manually verified in the Integrative Genomics Viewer (IGV).[45]

The candidate somatic mutations in tumor were included by over 1% VAF. And in TALs, somatic variant calls met the following criteria were included: (A) VAF>0.5%, (B) supported by at least 3 reads. All the candidate somatic mutations with synonymous or UTR variants were excluded. Germline variants from the matched control samples were subtracted from each tumor variant results.

RNA sequencing data analysis

RNA purity was checked using the kaiaoK5500Spectrophotometer (Kaiao, Beijing, China). RNA integrity and concentration were assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). According to the manufacturer's

protocol, mRNA libraries were prepared using the NEB Next Ultra RNA Library Prep Kit for Illumina (#E7530L, NEB, USA). The constructed RNA-seq libraries were paired-end sequenced on an Illumina NovaSeq 6000 sequencer (Illumina, San Diego, CA, USA), and 150 bp paired-end reads were generated. We removed the sequencing reads containing adaptor sequences and low-quality reads, which have too many Ns (>5%) and low-quality bases (>50% bases with quality <19), and obtained high-quality paired-end reads. Then these reads were aligned to the human genome (hg19) using HISAT2[46] (v2.1.0). Transcript assembly was performed using StringTie[47] (v1.2.3). Read count for each gene in each sample was counted by HTSeq[48] v0.6.0, and FPKM (Fragments Per Kilobase Million Mapped Reads) was then calculated to estimate the expression level of genes in each sample.

DESeq2[49] was used for differential gene expression analysis. The R package "clusterProfiler" was used for KEGG Pathway analysis. Single sample gene set enrichment analysis (ssGSEA)[16] was used to calculate the enrichment scores (ES) of immune cell types in the tumor microenvironment. Gene signatures of 28 immune cell types associated with innate and adaptive immunity were desposited in the Table S8.[50] We further subclassified the patients into immune groups using the Euclidean distance method and 'ward.D2' clustering method. We constructed the tumor-infiltrating lymphocytes (TIL) enrichment profiles by Pre-ranked GSEA method.[51] For each patient, expression levels of genes were *Z* score normalized across all patients and ranked in descending order according to the z-scores (or mean of z-scores for each group). Genes associated with immune signatures were compared to the above ranking utilizing the GSEA methods. The normalized enrichment score (NES) was obtained for each patient (or cohort). We identified the immune signature gene lists with false discovery rate (q-value) < 10% and NES>0 as enrichment, while false discovery rate<10% and NES<0 as depletion.

HLA typing and loss of heterozygosity

HLA type was determined using Optitype.[52] The polymorphisms in exons 2 and 3 of Class-I genes (HLA-A, HLA-B, HLA-C) were considered for HLA typing. Loss of heterozygosity (LOH) of HLA was assessed by LOH HLA.[20]

DSP data generation and analysis

In situ protein profiling was performed using GeoMx DSP Platform. Briefly, tissue slides were stained with a multiplexed panel of protein antibodies that contained a photocleavable indexing oligonucleotide, enabling subsequent readouts. Released indexing oligonucleotides from each region of interest (ROI) were collected and deposited into designated wells on a microtiter plate, allowing for well indexing of each ROI during nCounter (nCounter MAX system version 4.1.0.1) readout (direct protein hybridization). LUSC tissues from 42 patients were fixed with formalin and embedded with paraffin. Based on the H&E staining results, 3 separate regions (tumor-enriched region, tumor and normal cell-mixed region and normal cell-enriched regions), were taken from each individual tissue sample to make a tissue microarray for the study. The tissue microarray was stained with alpha-PanCK and alpha-CD45 antibody, then nuclei was counter stained with Syto13. ROIs with enriched PanCK and CD45 expression were selected separately in tumor-enriched region cores as well as in tumor and normal cell-mixed region cores. Meanwhile, ROIs with either CD45 high expression or CD45 low expression were selected in normal cell-enriched region cores. Spatially resolved multiplex protein profiling of selected ROIs was performed with GeoMx DSP platform and quantitated with ncounter. For data analysis, the raw counts were first normalized to External RNA Control Consortium (ERCC) spike-in control, followed by housekeeping normalization using Histone H3, GAPDH and S3 as the housekeeping protein. The normalized dataset was used for downstream data analysis such as differential target expression, heatmap and boxplot.

To explore the immune contexture in tumors, tumor edges and TALs, we evaluated the expression of immune-associated proteins (B2M, CD11c, CD20, CD3, CD4, CD56, CD45, CD68, CTLA4, GZMB, PD-1, PD-L1, HLA-DR) and compared the differences in three tissues by Kruskal-Wallis test and Wilcoxon test. Immune heterogeneity across ROIs was assessed by the range of pairwise Cosin similarity. Specifically, pairwise score in one ROI was defined as NA. If only two ROIs were included in one tissue, the ROI heterogeneity was assessed by the Cosin distance.

QUANTIFICATION AND STATISTICAL ANALYSIS

Comparisons of proportions and variables between different groups were performed with the Kruskal-Wallis test, and paired or unpaired Mann-Whitney-Wilcoxon test, as appropriate. Pearson correlation was used to analyze the correlation between univariates. All statistical analyses were performed in the R statistical environment version 3.3.4. Statistical significance was defined as a two-sided p value of <0.05. Survival curves were estimated with the Kaplan-Meier product-limit method and compared by log rank test.