Immune cell landscape analysis reveals prognostic immune cells and its potential mechanism in squamous cell lung carcinoma

Yongyong Wang and Jianji Guo

Cardio-Thoracic Surgery, First Affiliated Hospital of Guangxi Medical University, Nanning, China

ABSTRACT

Background. Squamous cell lung carcinoma (LUSC) was closely associated with smoking which was known to have a distant immunosuppression effect. In this study, we aimed to explore the relationship between immune cells and clinical outcomes of LUSC patients with smoking history.

Methods. The immune cell infiltration and RNA expression profiles of LUSC patients were collected from The Cancer Genome Atlas (TCGA). Then, the correlation between immune cell infiltration and clinical characteristics was explored. According to the level of immune cell infiltration, LUSC patients with smoking history were divided into high or low group to screen the differentially expressed lncRNAs and mRNAs. The prediction of target genes was performed by miRanda. Finally, the prognostic value of a certain signature was confirmed in an independent dataset.

Results. Higher abundance of tumor-infiltrating T follicular helper (Tfh) cells together with a lower abundance of resting memory CD4 T cells had been found in LUSC current reformed smokers for ≤15 years and current smoking patients. Moreover, Tfh cell infiltration was not only associated with better overall survival (OS) but also varied from different degrees of TNM stage. Low expression of lncRNA PWRN1 and its potential regulating genes DMRTB1, PIRT, APOBEC1, and ZPBP2 were associated with better OS. Combining PWRN1 and four regulating genes as a signature, patients with higher-level expression of the signature had shorter survival time in not only the TCGA but also in the GEO dataset.

Conclusions. It was found that Tfh cells presented higher infiltration in LUSC current reformed smokers for ≤15 years and current smokers, while resting memory CD4 T cells had lower infiltration. The signature consisting of PWRN1 as well as its predicted targeted mRNAs was dysregulated in different levels of Tfh cell infiltration and might indicate patients' OS.

Subjects Bioinformatics, Gastroenterology and Hepatology, Oncology, Respiratory Medicine
Keywords Squamous cell lung carcinoma, T follicular helper cells, PWRN1, ceRNA, Overall survival

INTRODUCTION

Since 2010, lung cancer has been the leading cause of cancer-related death worldwide (*Ding et al., 2019*). Non-small cell lung cancer (NSCLC), the most common kind of lung cancer

(accounting for 83% of total lung cancer cases), has a poor 5-year survival rate, which is lower than 20%, although great efforts and progresses have been made in chemotherapy, radiotherapy, surgery and immunotherapy (*Miller et al., 2019*; *Wang et al., 2020*).

For decades, the risk of lung cancer from smoking has been well established through epidemiological methods (*Alberg et al., 2007*; *Okamoto et al., 2014*). Squamous cell lung carcinoma (LUSC), a subtype of NSCLC, had a tight association with tobacco smoking as 96% of LUSC cases had smoking history (*Cancer Genome Atlas Research, 2012*; *Kenfield et al., 2008*; *Khuder & Mutgi, 2001*). Tobacco smoking promotes cancer progression mainly in two ways. Firstly, cigarette smoking is tumorigenic by causing gene alterations, such as the major tumor suppressor gene TP53, despite only a small proportion of smokers eventually developed into cancer (*Alberg, Brock & Samet, 2005*). Secondly, it could influence defense mechanism by suppressing NK cell number and activity as well as cytotoxic T cell activity (*Chalmer, Holt & Keast, 1975*; *Mio et al., 1997*; *Ng et al., 2006*; *Sopori, 2002*; *Stampfli & Anderson, 2009*; *Swann et al., 2007*; *Woodruff et al., 2005*), being a negative prognostic factor for cancer patients. However, few studies have learned about the relationship among changes in smoking status, mortality, and immune cell infiltration, which was focused in our research.

The T follicular helper (Tfh) cell, a novel type of CD4+ T helper cell that has been discovered recently, interacts closely with B cells resulting in B cell proliferation and activation (*Ramiscal & Vinuesa, 2013*). The effect of tobacco smoking on Tfh cells infiltration in LUSC patients has not been studied thoroughly in the past investigations. Previous studies found that Tfh cell was positively associated with OS of NSCLC (*Ma et al., 2016*). However, the underlying mechanism still need further exploration and in this study, we tried to explain it by the dysregulated long non-coding RNA (lncRNA)-mRNA pairs in patients with different degrees of immune cell infiltration. That has not been reported before, so present research may offer some novel points to investigate Tfh cells.

LncRNAs, transcripts longer than 200 bp without protein-coding potential, have been found to play an important role in tumor formation and development (*Heery et al., 2017*; *Lin & Yang, 2018*). LncRNAs could regulate the expression of mRNA by competing with shared miRNAs, which was defined as competing endogenous RNAs (ceRNAs) (*Li et al., 2019*). CeRNA network had been investigated in multiply researches to probe prognosis biomarker of cancer or explain molecular processes during tumorigenesis and metastasis (*Cheng et al., 2019*; *Li et al., 2020*). Combined with OS and differentially expressed analysis, the related research of lncRNAs will also help to reveal some novel prognosis biomarkers.

In this study, the immune cell infiltration data and RNA expression profiles of LUSC patients were collected from The Cancer Genome Atlas (TCGA). Then, the correlation of immune cell infiltration with patient clinical characteristics such as gender, TNM stage, and smoking history, were explored. According to the different abundance of immune cell infiltration that presented in reformed smokers, current smokers, and non-smoking patients, LUSC patients with smoking history were divided into high and low infiltration group. The differentially expressed lncRNAs and mRNAs were screened to construct ceRNA network so that we can further find explanations of immune cell in prognosis

evaluation and potential prognosis biomarkers. Finally, the prognostic value of certain signature was confirmed in another dataset downloaded from GEO.

MATERIALS AND METHODS

Data collection

Firstly, the immune cell fraction and survival data of 490 LUSC patients from TCGA were downloaded, which had been reported by *Thorsson et al. (2018)*. The relative fraction of 22 immune cell types of each patient was estimated using CIBERSORT (*Newman et al., 2015*) based on TCGA RNA-seq data. CIBERSORT is a computational method that can accurately calculate relative levels of various immune cell types in complex gene expression mixtures. Using gene expression signatures of 547 genes (LM22 files) as an input matrix, CIBERSORT can characterize and quantify immune cell subtypes, such as Tfh cells and resting memory T cells. With TCGA LUSC RNA-seq expression data and LM22 files being input files, CIBERSORT was implemented in ''relative mode'' to estimate the relative abundance of tumor infiltrating immune cells.

Secondly, the clinical information, including age, gender, TNM stage, and smoking history (Table S1), together with TCGA RSEM normalized RNA-seq gene expression data (Illumina HiSeq) of 490 LUSC patients were obtained from UCSC Xena (https://xena.ucsc.edu/public). TNM stage is a kind of cancer staging system developed by the American Joint Committee on Cancer. T, N, and M mean tumor, node and metastasis, respectively. Lifelong non-smokers ($n = 17$) means person who was not smoking at the time of the interview and has smoked less than 100 cigarettes in their life. Current reformed smokers were defined as people who was not smoking at the time of the interview but has smoked at least 100 cigarettes in their life. According to the duration of cessation, they were classified into current reformed smokers for >15 years ($n = 78$) and ≤ 15 years ($n = 247$). The number of current smokers was 132. Given that people haven't smoked for 15 years could discontinue lung cancer screening (*Moyer & Force, 2014*) and the sample size of non-smokers was too small, we combined non-smokers and current reformed smokers for >15 years as ''non- + ex-smokers'' group to do the comparison. Finally, the validation cohort (GSE50081, Affymetrix HG-U133_Plus_2) was downloaded from GEO (Gene Expression Omnibus, https://www.ncbi.nlm.nih.gov/gds/) to confirm the prognosis value of certain key signature. Detailed workflow was shown in Fig. 1.

Overall survival (OS) analysis

On the basis of immune cell fraction, LUSC patients were divided into two groups. Combined with patients' survival data, tumor infiltrating immune cells that had prognosis value were screened out. OS analysis was performed via cox proportional hazards regression model by R language with survminer package, and visualized by survival package. After inputting the survival time and the end point information (means patient dead or alive), Kaplan–Meier survival analysis and log-rank tests were used to obtain the survival curve. In OS analysis, the cutoff of high and low group was set by the optimal *P* value according to the infiltration abundance of each patient.
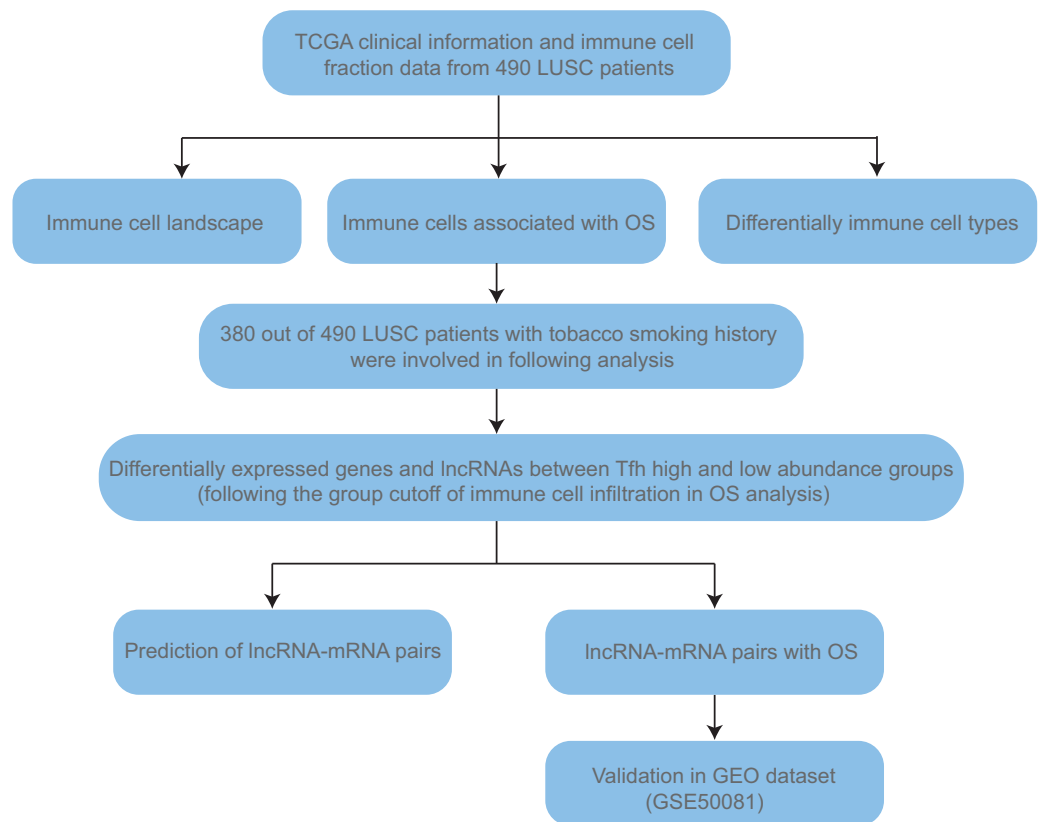
Figure 1 The workflow of this study.

Full-size ⬛ DOI: 10.7717/peerj.9996/fig-1

Differentially expressed genes (DEGs) and lncRNAs (DElncRNAs)

380 out of 490 LUSC patients with smoking history were involved in this analysis. We continued to use the same cutoff with OS analysis to divide high and low group following the infiltration abundance estimated by CIBERSORT of each patient. According to immune cell fraction, patients were classified into high and low expression groups. The fold change expression of each gene in high and low groups was calculated and log2-transformed. The DEGs and DElncRNAs between two groups were screened with the threshold of |log2 (fold change)| > 1 and adjusted P value < 0.05 (P value was adjusted by FDR method). The OS analysis of DEGs and DElncRNAs was carried out using similar methods as described in OS analysis section.

Prediction of lncRNA-mRNA pair and ceRNA network construction

Based on the targeted miRNA dataset of lncRNA or mRNA downloaded from the miRanda (http://www.microrna.org/microrna/home.do), DEGs and DElncRNA target miRNAs were found. After inputting a two-column file including the information of DElncRNAs and its target miRNAs, Cytoscape (venison 3.6.1) would exhibit a ceRNA network.

Statistical analysis

All statistical analysis in this study was performed using R language. Two group tests were carried out with un-paired t-test or Wilcoxon test, and multiple group comparison was carried out with kruskal-wallis method.

RESULT

The correlation between immune cell infiltration and clinical characteristics of LUSC patients

We investigated the relationship between the change of immune cell infiltration and clinical characteristics, such as TNM stage, gender and tobacco smoking history. Firstly, the abundance of 22 tumor infiltrating immune cells in 490 patients was analyzed by CIBERSORT. The abundances of M2 Macrophages, CD8 T cells, plasma cells, resting memory CD4 T cells, and Tfh cells in LUSC tumor were higher than that of most other immune cells (Fig. 2A). Secondly, the status of tumor infiltrating immune cell was compared between female and male in this cohort. Compared to male LUSC patients, female LUSC patients had more M1 macrophages, resting dendritic cells, and activated memory T cells CD4 in the tumor ($P < 0.05$, Fig. 2B). Furthermore, the immune cell infiltration at various stages was also analyzed, and only Tfh cell infiltration exhibited statistical difference at various TNM stage. Compared to TNM stage I, Tfh cell infiltration level decreased at TNM stage II ($P = 0.049$, Fig. 2C), while the infiltration of Tfh cell slightly increased at TNM stage III (Fig. 2C). Tfh cell infiltration in detailed TNM stage was compared for further exploration. With tumor growing, the infiltration of Tfh cells declined in T2 and T3 stage, compared to T1 stage ($P = 0.021$ and 0.002, Fig. 2D). However, in T4 stage, the abundance of Tfh cells showed a slight increase (Fig. 2D). In different degrees of lymphatic node metastasis, Tfh cells fraction in N2 stage exhibited significantly rising, compared with N0 stage ($P = 0.032$, Fig. 2E). Moreover, there were other immune cell infiltrations showing significant differences in detailed TNM stage. For example, Macrophages M0 infiltration increased along with the rising of tumor size in T1, T2, and T3, but not in T4 (Fig. S1A). The infiltration of activated mast cells and neutrophils was higher in T2, while the infiltration of activated NK cells presented opposed trend in T2, compared to T1 stage (Fig. S1B & Figs. 2D & 2E). Meanwhile, neutrophils also showed higher infiltration in M1 (Fig. S1). The infiltration of resting mast cells and CD8 T cells declined in T3 (Fig. S1C & Fig. 2F). The infiltration of M1 macrophages and T cells CD8 significantly increased in N1 and N2 stage, respectively, compared with N0 stage (Figs. S1G–S1H).

DEGs, DElncRNAs and construction of ceRNA network

In order to find immune cells that might be affected by tobacco smoking in LUSC, the abundance of 22 tumor infiltrating immune cells were assessed by CIBERSORT. Two types of immune cells exhibited significant difference among non- and ex-smokers, current reformed smokers for ≤ 15 years and current smokers. Resting memory CD4 T cell infiltration in patients not smoking for ≤ 15 years ($P = 0.012$) and current smokers decreased ($P = 0.001$, Fig. 3A), while Tfh cell infiltration in the two groups both increased ($P = 0.03, 0.022$, Fig. 3B), compared to non- + ex-smokers. However, only Tfh cells fraction
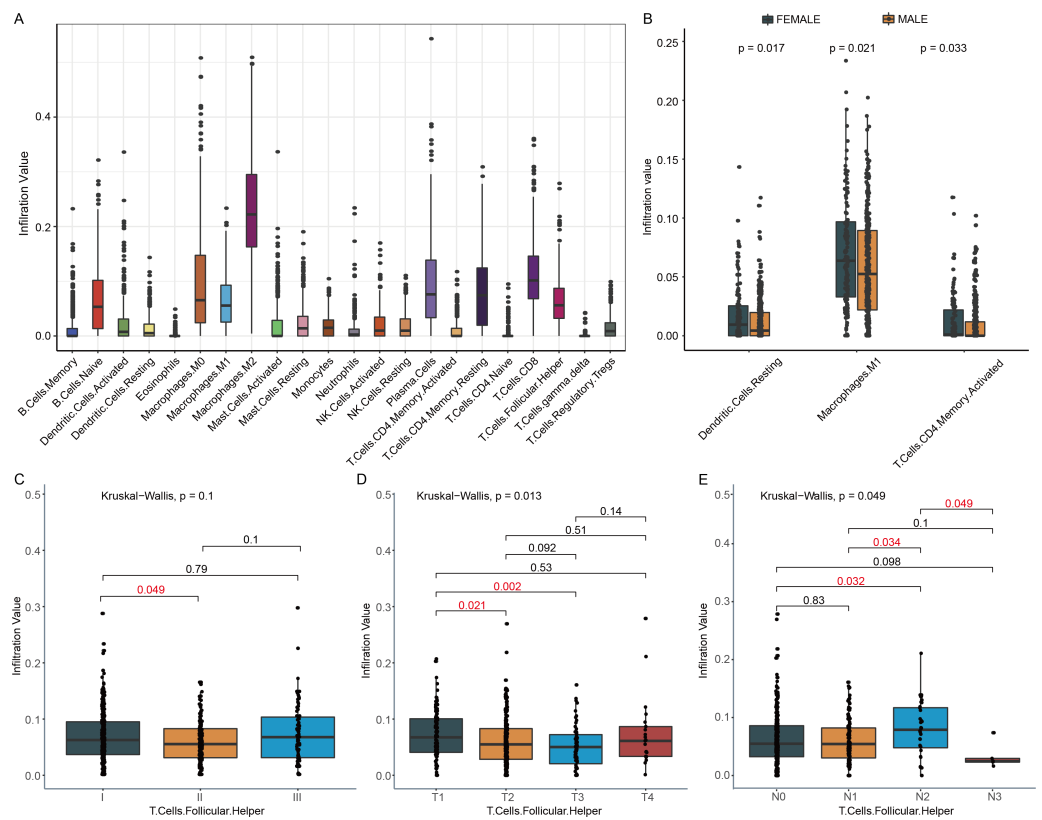
Figure 2 Immune cell infiltration and clinical characters of LUSC patients. (A) Twenty-two immune cells infiltration landscape in LUSC patients was observed. (B) This boxplot showed the distribution of three immune cells in females and males. (C) Tfh cell infiltration in different TNM stages was exhibited. (D) In the T2 and T3 stage, Tfh cells had lower infiltration, compared to the T1 stage. T, tumor (E) In the N2 stage, Tfh cells infiltration had significantly increased, while in the N3 stage, the trends turned into the opposite. N, regional lymph nodes.

Full-size 🖼 DOI: 10.7717/peerj.9996/fig-2

was related to OS with $P = 0.016$ and hazard ratio (HR) = 0.71 (Fig. 3C). Hence, in the following investigation, 380 LUSC patients with smoking history were divided into high and low group according to the infiltration of Tfh cells to identify functional genes that related to clinical outcomes.

In the OS analysis, the cutoff threshold of high and low group was set by the best P value. And in the analysis of DEGs and DElncRNAs, the group cutoff value was set in accordance with that. From differentially expression analysis, 61 DEGs and 2 DElncRNAs were screened out with the threshold of $|\log2$ (fold change) $> 1|$ and adjusted $P < 0.05$ and showed in the volcano plot (Fig. 4A). The prediction of lncRNA-miRNA and miRNA-mRNA functional pairs were both carried out on miRanda. A ceRNA network centered on DElncRNA PWRN1 was built and visualized in Fig. 4B. The red rhombus presented up-regulated DElncRNA PWRN1, and the yellow triangle pointed to up-regulated DEGs. Blue circular was on behalf of miRNA that PWRN1 and DEGs could both combine. In Fig. 4C, the binding site of partial paired DElncRNA-miRNA-DEGs was shown and all of them had a mirSVR score <-0.1 and PhastCons score >0.5. mirSVR and PhastCons score indicate
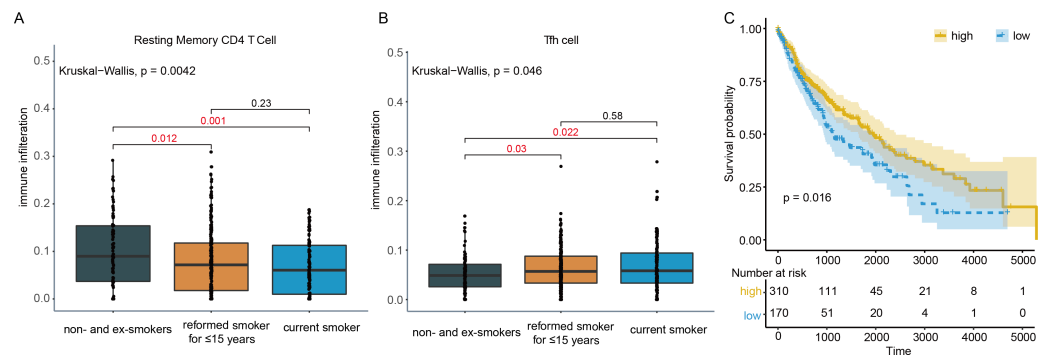
Figure 3 **Immune cell infiltration in smoking and non-smoking LUSC patients.** (A) Resting memory CD4 T cells showed lower infiltration in current reformed smokers for $\leq$ 15 years and current smokers. (B) In current reformed smokers for $\leq$ 15 years and current smokers, Tfh cell infiltration increased significantly. (C) Higher Tfh cell infiltration patients had better OS.

Full-size 🖼 DOI: 10.7717/peerj.9996/fig-3

thermodynamic stability and conservation, respectively. Lower mirSVR score presents better binding stability and higher PhastCons score means more conservative

OS analysis and validation

To evaluate the prognostic value of DEGs and PWRN1 in ceRNA network, OS analysis was performed. Comparing to high expression group, the group with low expression of PWRN1 ($P = 0.021$), DMRTB1 ($P = 0.014$), PIRT ($P = 0.013$), APOBEC1 ($P = 0.025$), and ZPBP2 ($P = 0.0039$) had better OS (Figs. 5A–5E). HR value of high PWRN1, DMRTB1, PIRT, APOBEC1, and ZPBP2 expression group was 1.74, 1.75, 1.72, 1.58, 1.46, and 1.55 respectively (Table 1). The 95% CI of them was also exhibited in Table 1. The cutoff value of high and low expression in this OS analysis was also set by the best P value. When we averaged the expression of PWRN1, DMRTB1, PIRT, APOBEC1, and ZPBP2 to make them as a signature and divided the patients by median value of signature expression, low expression group showed better OS than high expression group (Fig. 5F, $P = 0.0058$, HR = 1.55). Subsequently, the finding was validated in GSE50081. The above-mentioned five genes were combined and 34 LUSC patients with smoking history were grouped in the same way. The low signature expression group also demonstrated better prognosis clinical outcomes with a P value of 0.046 (Fig. 6) and a HR score of 3.57 (Table 1).

DISCUSSION

NSCLC is the most common kind of lung cancer (accounting for 83% of cases) with a poor 5-year survival of less than 20%. LUSC, a subtype of NSCLC, associated closely to tobacco smoking. Cigarette smoking exposure could not only induce genetic alteration, but also suppress the immune system so that to contribute to carcinogenesis and metastasis. Previous studies found that the infiltration of some immune cells, such as Tfh cells, was associated with better OS. Therefore, in this study, we tried to explore the reason why
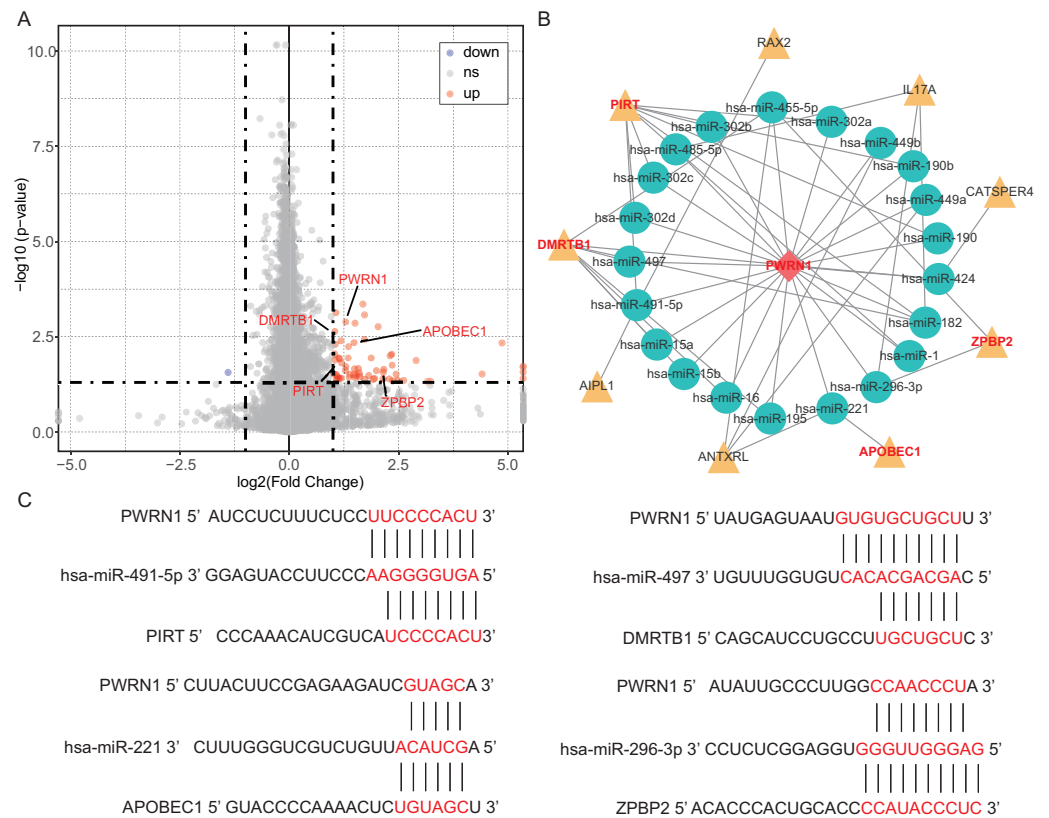
Figure 4 DEGs, DElncRNAs and construction of ceRNA network. (A) The volcano plot showed the DEGs and DElncRNAs. The red point indicated DEGs or DElncRNAs were up-regulated in high Tfh cells infiltration group, while the blue point implied they were down-regulated. ns, not significant. (B) The ceRNA network was built via Cytoscape. Yellow triangle, up-regulated genes. (C) The binding site of lncRNA-miRNA-mRNA was predicted by miRanda.

Full-size 🖼 DOI: 10.7717/peerj.9996/fig-4

higher infiltration of Tfh cells corresponds to better clinical outcomes from the perspective of ceRNA.

Firstly, the immune cell infiltrations in LUSC patients were observed. And we found that M2 Macrophage, CD8 T cells, plasma cells, resting memory CD4 T cells, and Tfh cells had higher infiltration compared to other immune cells. In NSCLC, M2 macrophage promotes epithelial to mesenchymal transition (EMT) and causes malignancy (*Guo et al., 2019*). Higher infiltration of CD8 T cells, plasma cells and Tfh cells was associated with better OS of patients with cancer (*Knief et al., 2016*; *Miotto et al., 2010*; *Zhu & Zhang, 2019*). Subsequently, we analyzed the infiltration of immune cells in different TNM stage, and only Tfh cells presented significant differences in various TNM stages. Compared to stage I, Tfh cell exhibited a slight decline in stage II, but in stage III, the trend was not going on, which was partly in agreement with previous report (*Guo et al., 2017*). In the detailed TNM stage, Figs. 2D and 2E actually showed negative correlation to T and positive association to N stage. There are two possible explanations about the decline in stage II. First, for the immune infiltration abundance of each cell estimated by CIBERSORE is relative value,
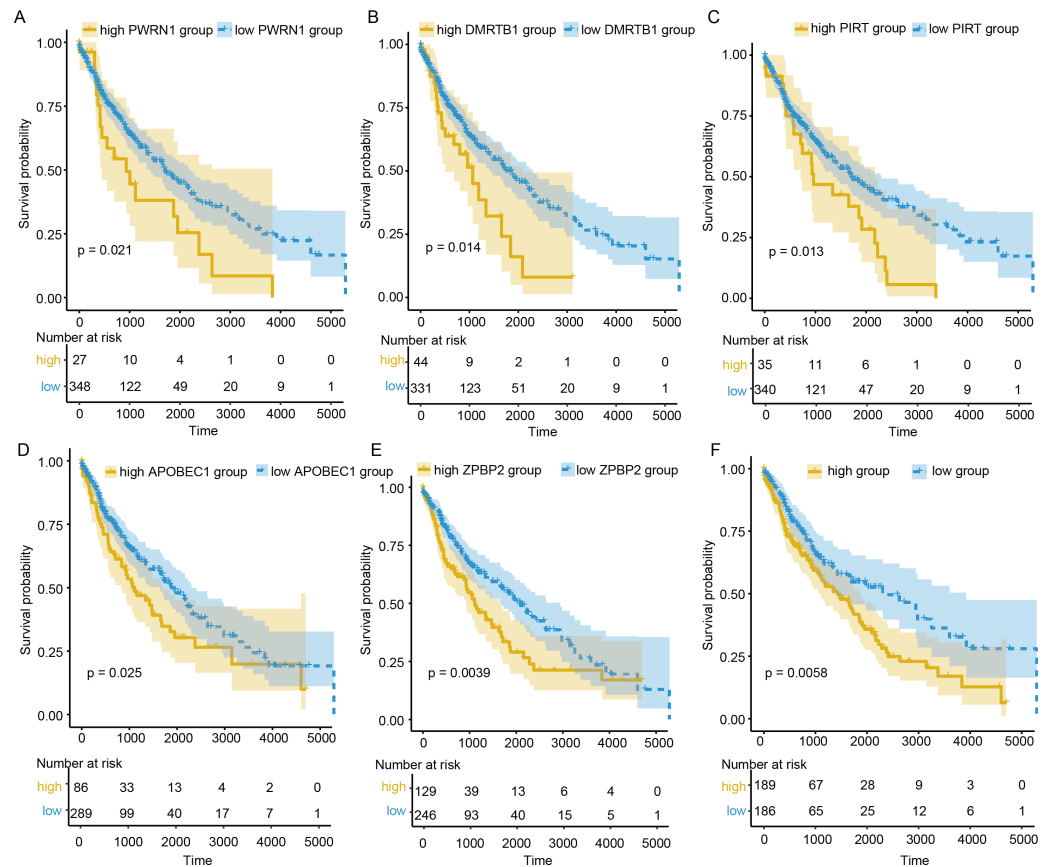
Figure 5 OS analysis of DElncRNAs, DEGs and signature. (A–E) The patients with low expression PWRN1 (A), DMRTB1 (B), PIRT (C), APOBEC1 (D) and ZPBP2 (E), had better OS. (F) Low signature group showed longer survival time.

Full-size ☑ DOI: 10.7717/peerj.9996/fig-5

the decline of Tfh cell infiltration abundance in stage II may result from the increase or decrease of other immune cells. Second, in different stages of NSCLC, the change of Tfh subtypes is inconsistent. The Tfh1 cell abundance in stage II-IV decreased, while Tfh2 and Tfh17 abundances in stage II-IV increased (*Qiu et al., 2018*) and the ratio of Tfh1, Tfh2, and Tfh17 abundances might lead to the Tfh cell infiltration abundance changing. Further, benefit from immune escape, tumor cells can migrate to other space. CD4+CXCR5+ T cell population was examined to define Tfh cells (*Ma et al., 2016*) and PD-1+ CD4+ CXCR5+ T cells presented lower capacity to assist CD8 T cell (*Ma et al., 2016*). Hence, we supposed that this Tfh cell subtypes might be chosen higher infiltration so that the tumor stage and metastasis could progress. The decrease of Tfh cell abundance in N3 stage might because the sample size of N3 stage patients was too small to make result not reliable (the patient number in N1, N2, and N3 stage was 304, 128, 36, and 5). Further, previous reports showed that female patients with NSCLC had better OS (*Freudenstein et al., 2020*). In our results, we found some clues to explain it. The dendritic cells (*Freudenstein et al., 2020*), M1 macrophages (*Cornelissen et al., 2014*), and activated CD4 T memory cells (*Zhang et al., 2020*) that exhibited anti-tumor capacities were all demonstrated higher infiltration in

Table 1 Hazard Ratio and *P* value of each group.

Source	Group	Hazard Ratio (HR)	logrank P	lower 95% CI	upper 95% CI
TCGA	High PWRN1 group	1.74	0.02	1.08	2.81
TCGA	High DMRTB1 group	1.75	0.01	1.11	2.76
TCGA	High PIRT group	1.72	0.01	1.11	2.66
TCGA	High ZPBP2 group	1.58	0.004	1.15	2.16
TCGA	High APOBEC1 group	1.46	0.026	1.05	2.05
TCGA	High signature group	1.55	0.006	1.13	2.12
GSE50081	High signature group	3.57	0.046	0.94	13.59

females, compared to the male. That might because the different levels of estrogen and androgen in females and males. The differential effect of estrogen on immune function reflects not only the concentration of estrogen, but also reflects the density, distribution and type of estrogen in immune cells (*Klein & Flanagan, 2016*). Estrogen receptor (ERs) are expressed in various immune cells, such as macrophages, T cells and dendritic cells (*Phiel et al., 2005*). Estrogen could influence functional activity of innate immune cells and adaptive immune responses. Androgens exhibited higher concentrations in post-pubertal males than females could inhibit immune cell activity (*Roberts, Walker & Alexander, 2001*).

Given that tobacco smoking exposure was a vital factor of LUSC carcinogenesis, the immune cell infiltration in LUSC current reformed smokers, current smokers and non-smoker patients was analyzed. The infiltration of Tfh cells increased in smokers, while that of resting memory CD4 T cells decreased. Tobacco smoking exposure could result in CD4+ memory T cells switching from resting to activation (*Li et al., 2018*). Further, cigarette smoking could also trigger the activation of dendritic cells to facilitate CD4+ T cells differentiation into Tfh cells (*Qiu et al., 2017*; *Ramiscal & Vinuesa, 2013*). That might be the explanation of increased infiltration of Tfh cells and decreased infiltration of resting memory CD4 T cells. After cox proportional hazards regression analysis, patients with high infiltration of Tfh cells showed better clinical outcomes, which was also consistent with previous findings (*Ma et al., 2016*; *Xu et al., 2020*).

To find potential molecule mechanism between Tfh cells and OS, LUSC patients with smoking history were divided into two groups according to the infiltration of Tfh cells so that DElncRNAs and DEGs could be screened out. It was found that the combination of DElncRNA PWRN1, DEG DMRTB1, PIRT, APOBEC1, and ZPBP2 was significantly related to OS. However, the function of these genes have not been studied adequately in LUSC. Prader-Willi region non-protein coding RNA 1(PWRN1) located in chromosome 15 and up-regulated PWRN1 suppressed cell proliferation and tumor growth in gastric cancer (*Chen et al., 2018*). But in our study, high expression PWRN1 was related to worse clinical outcomes, which might be because PWRN1 was a strongly tissue-specific lncRNA so that PWRN1 could both act as proliferation-promoting and tumor suppressor role. For example, NEUROD1/2 functions as growth-promoting gene in HPNE cells, but as suppressor of proliferation in HMEC cells (*Sack et al., 2018*). DMRT like family B with proline rich C-terminal 1 (DMRTB1) mutation might relate to human colorectal cancer
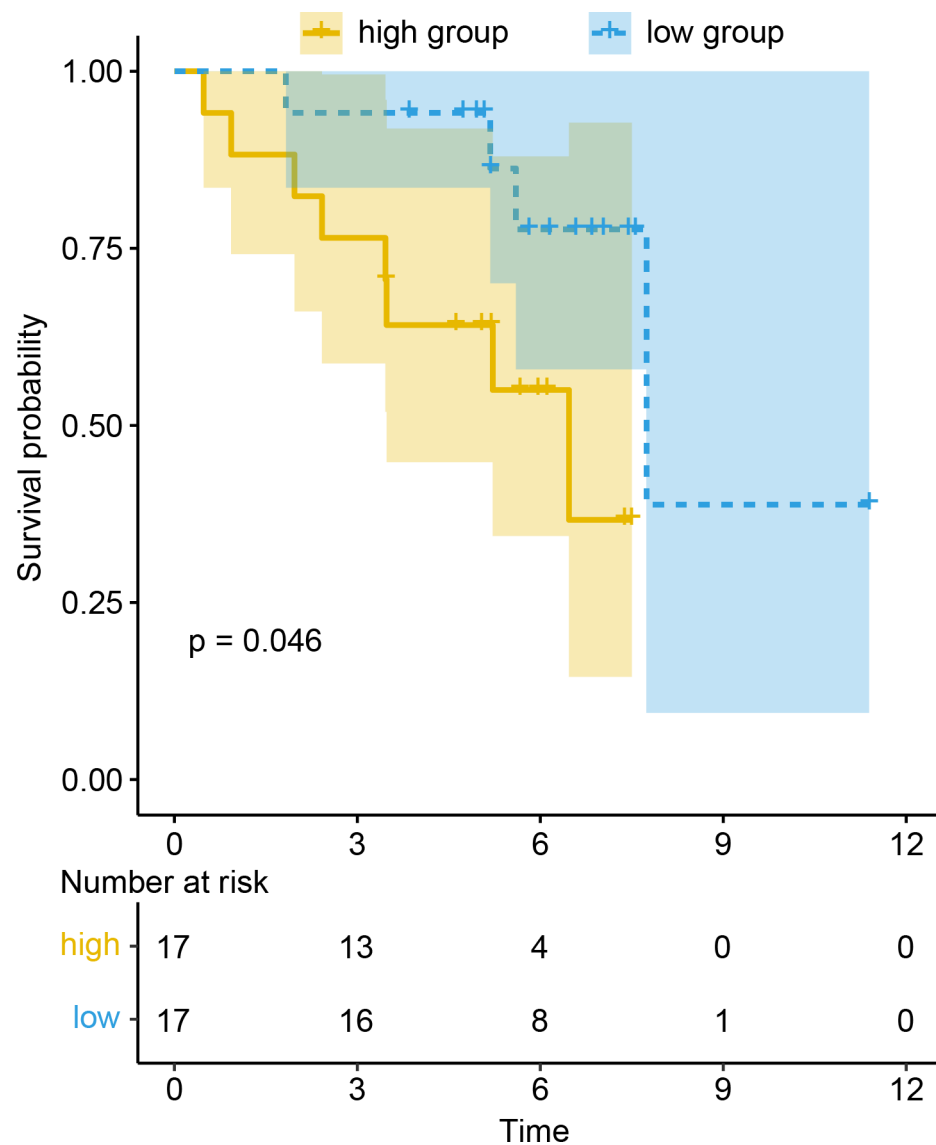
Figure 6 **Prognosis value of signature was confirmed in the GEO dataset.**
Full-size ◩ DOI: 10.7717/peerj.9996/fig-6

liver metastasis (*Ma et al., 2018*). Ahmadreza Niavarani et al. found that strong expression of apolipoprotein B mRNA editing enzyme catalytic subunit 1 (APOBEC1) was associated with short survival time in pan-cancer, which was in line with our result (*Niavarani et al., 2018*). And that might because APOBEC1 behave as a consistently hypo-methylated gene in pan-cancer (*Niavarani et al., 2018*). Wiemels JL et al. identified new risk loci, located on 17q12 near Zona pellucida binding protein 2 (ZPBP2), could regulate gene expression by local interactions with ZPBP2, impacting function of hematopoietic and growth-regulation pathways in childhood acute lymphoblastic leukemia (*Wiemels et al., 2018*). The DElncRNA or DEGs in PWRN1-mediated ceRNA network was related to clinical outcomes or cancer

development somehow. Therefore, I reckoned that the PWRN1-mediated ceRNA network was a potential prognostic biomarker in LUSC.

CONCLUSIONS

In conclusion, it was found that Tfh cells presented higher infiltration in LUSC current reformed smokers for ≤ 15 years and current smokers, while resting memory CD4 T cells had lower infiltration. Tfh cells infiltration was not only associated with better OS, but also varied from different degrees of the TNM stage. The signature consisting of PWRN1 and its predicted targeted mRNAs was dysregulated in different degrees of Tfh cell infiltration and might indicate patients' OS. The prognostic value of the signature was also confirmed in GEO dataset (GSE50081).

ADDITIONAL INFORMATION AND DECLARATIONS

REFERENCES

Alberg AJ, Brock MV, Samet JM. 2005. Epidemiology of lung cancer: looking to the future. *Journal of Clinical Oncology* **23**:3175–3185 DOI 10.1200/JCO.2005.10.462.

Alberg AJ, Ford JG, Samet JM, American College of Chest P. 2007. Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* **132**:29S–55S DOI 10.1378/chest.07-1347.

Cancer Genome Atlas Research N. 2012. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**:519–525 DOI 10.1038/nature11404.

Chalmer J, Holt PG, Keast D. 1975. Cell-mediated immune responses to transplanted tumors in mice chronically exposed to cigarette smoke. *Journal of the National Cancer Institute* **55**:1129–1134 DOI 10.1093/jnci/55.5.1129.

Chen Z, Ju H, Yu S, Zhao T, Jing X, Li P, Jia J, Li N, Tan B, Li Y. 2018. Prader-Willi region non-protein coding RNA 1 suppressed gastric cancer growth as a competing endogenous RNA of miR-425-5p. *Clinical Science* **132**:1003–1019 DOI 10.1042/CS20171588.

Cheng Z, Yu C, Cui S, Wang H, Jin H, Wang C, Li B, Qin M, Yang C, He J, Zuo Q, Wang S, Liu J, Ye W, Lv Y, Zhao F, Yao M, Jiang L, Qin W. 2019. circTP63 functions as a ceRNA to promote lung squamous cell carcinoma progression by upregulating FOXM1. *Nature Communications* **10**:Article 3200 DOI 10.1038/s41467-019-11162-4.

Cornelissen R, Lievense LA, Maat AP, Hendriks RW, Hoogsteden HC, Bogers AJ, Hegmans JP, Aerts JG. 2014. Ratio of intratumoral macrophage phenotypes is a prognostic factor in epithelioid malignant pleural mesothelioma. *PLOS ONE* **9**:e106742 DOI 10.1371/journal.pone.0106742.

Ding M, Li F, Wang B, Chi G, Liu H. 2019. A comprehensive analysis of WGCNA and serum metabolomics manifests the lung cancer-associated disordered glucose metabolism. *Journal of Cellular Biochemistry* **120**:10855–10863 DOI 10.1002/jcb.28377.

Freudenstein D, Litchfield C, Caramia F, Wright G, Solomon BJ, Ball D, Keam SP, Neeson P, Haupt Y, Haupt S. 2020. TP53 status, patient sex, and the immune response as determinants of lung cancer patient survival. *Cancers* **12(6)**:1535 DOI 10.3390/cancers12061535.

Guo Z, Liang H, Xu Y, Liu L, Ren X, Zhang S, Wei S, Xu P. 2017. The role of circulating T follicular helper cells and regulatory cells in non-small cell lung cancer patients. *Scandinavian Journal of Immunology* **86**:107–112 DOI 10.1111/sji.12566.

Guo Z, Song J, Hao J, Zhao H, Du X, Li E, Kuang Y, Yang F, Wang W, Deng J, Wang Q. 2019. M2 macrophages promote NSCLC metastasis by upregulating CRYAB. *Cell Death & Disease* **10**:Article 377 DOI 10.1038/s41419-019-1618-x.

Heery R, Finn SP, Cuffe S, Gray SG. 2017. Long non-coding RNAs: key regulators of epithelial-mesenchymal transition, tumour drug resistance and cancer stem cells. *Cancers* **9(4)**:38 DOI 10.3390/cancers9040038.

Kenfield SA, Wei EK, Stampfer MJ, Rosner BA, Colditz GA. 2008. Comparison of aspects of smoking among the four histological types of lung cancer. *Tobacco Control* **17**:198–204 DOI 10.1136/tc.2007.022582.

Khuder SA, Mutgi AB. 2001. Effect of smoking cessation on major histologic types of lung cancer. *Chest* **120**:1577–1583 DOI 10.1378/chest.120.5.1577.

Klein SL, Flanagan KL. 2016. Sex differences in immune responses. *Nature Reviews Immunology* **16**:626–638 DOI 10.1038/nri.2016.90.

Knief J, Reddemann K, Petrova E, Herhahn T, Wellner U, Thorns C. 2016. High density of tumor-infiltrating B-lymphocytes and plasma cells signifies prolonged overall survival in adenocarcinoma of the esophagogastric junction. *Anticancer Research* **36**:5339–5345 DOI 10.21873/anticanres.11107.

Li R, Yang YE, Jin J, Zhang MY, Liu X, Liu XX, Yin YH, Qu YQ. 2019. Identification of lncRNA biomarkers in lung squamous cell carcinoma using comprehensive analysis of lncRNA mediated ceRNA network. *Artificial Cells, Nanomedicine, and Biotechnology* **47**:3246–3258 DOI 10.1080/21691401.2019.1647225.

Li S, Cui Z, Zhao Y, Ma S, Sun Y, Li H, Gao M, Li N, Wang Y, Tong L, Song M, Yin Z. 2020. Candidate lncRNA-microRNA-mRNA networks in predicting non-small cell lung cancer and related prognosis analysis. *Journal of Cancer Research and Clinical Oncology* **146**:883–896 DOI 10.1007/s00432-020-03161-6.

Li X, Li J, Wu P, Zhou L, Lu B, Ying K, Chen E, Lu Y, Liu P. 2018. Smoker and non-smoker lung adenocarcinoma is characterized by distinct tumor immune microen-vironments. *Oncoimmunology* **7**:e1494677 DOI 10.1080/2162402X.2018.1494677.

Lin C, Yang L. 2018. Long noncoding RNA in cancer: wiring signaling circuitry. *Trends in Cell Biology* **28**:287–301 DOI 10.1016/j.tcb.2017.11.008.

Ma QY, Huang DY, Zhang HJ, Chen J, Miller W, Chen XF. 2016. Function of follicular helper T cell is impaired and correlates with survival time in non-small cell lung cancer. *International Immunopharmacology* **41**:1–7 DOI 10.1016/j.intimp.2016.10.014.

Ma YS, Huang T, Zhong XM, Zhang HW, Cong XL, Xu H, Lu GX, Yu F, Xue SB, Lv ZW, Fu D. 2018. Proteogenomic characterization and comprehensive integrative genomic analysis of human colorectal cancer liver metastasis. *Molecular Cancer* **17**:139 DOI 10.1186/s12943-018-0890-1.

Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL, Siegel RL. 2019. Cancer treatment and survivorship statistics, 2019. *CA: A Cancer Journal for Clinicians* **69**:363–385 DOI 10.3322/caac.21565.

Mio T, Romberger DJ, Thompson AB, Robbins RA, Heires A, Rennard SI. 1997. Cigarette smoke induces interleukin-8 release from human bronchial epithelial cells. *American Journal of Respiratory and Critical Care Medicine* **155**:1770–1776 DOI 10.1164/ajrccm.155.5.9154890.

Miotto D, Lo Cascio N, Stendardo M, Querzoli P, Pedriali M, De Rosa E, Fabbri LM, Mapp CE, Boschetto P. 2010. CD8+ T cells expressing IL-10 are asso-ciated with a favourable prognosis in lung cancer. *Lung Cancer* **69**:355–360 DOI 10.1016/j.lungcan.2009.12.012.

Moyer VA, Force USPST. 2014. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine* **160**:330–338 DOI 10.7326/M13-2771.

Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. 2015. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* **12**:453–457 DOI 10.1038/nmeth.3337.

Ng SP, Silverstone AE, Lai ZW, Zelikoff JT. 2006. Effects of prenatal exposure to cigarette smoke on offspring tumor susceptibility and associated immune mechanisms. *Toxicological Sciences* **89**:135–144 DOI 10.1093/toxsci/kfj006.

Niavarani A, Shahrabi Farahani A, Sharafkhah M, Rassoulzadegan M. 2018. Pan-cancer analysis identifies prognostic high-APOBEC1 expression level implicated in cancer in-frame insertions and deletions. *Carcinogenesis* **39**:327–335 DOI 10.1093/carcin/bgy005.

Okamoto T, Suzuki Y, Fujishita T, Kitahara H, Shimamatsu S, Kohno M, Morodomi Y, Kawano D, Maehara Y. 2014. The prognostic impact of the amount of tobacco smoking in non-small cell lung cancer–differences between adenocarcinoma and squamous cell carcinoma. *Lung Cancer* **85**:125–130 DOI 10.1016/j.lungcan.2014.06.006.

Phiel KL, Henderson RA, Adelman SJ, Elloso MM. 2005. Differential estrogen receptor gene expression in human peripheral blood mononuclear cell populations. *Immunology Letters* **97**:107–113 DOI 10.1016/j.imlet.2004.10.007.

Qiu L, Yu Q, Zhou Y, Zheng S, Tao J, Jiang Q, Yuan G. 2018. Functionally impaired follicular helper T cells induce regulatory B cells and CD14(+) human leukocyte antigen-DR(-) cell differentiation in non-small cell lung cancer. *Cancer Science* **109**:3751–3761 DOI 10.1111/cas.13836.

Qiu SL, Zhang H, Tang QY, Bai J, He ZY, Zhang JQ, Li MH, Deng JM, Liu GN, Zhong XN. 2017. Neutrophil extracellular traps induced by cigarette smoke activate plasmacytoid dendritic cells. *Thorax* **72**:1084–1093 DOI 10.1136/thoraxjnl-2016-209887.

Ramiscal RR, Vinuesa CG. 2013. T-cell subsets in the germinal center. *Immunological Reviews* **252**:146–155 DOI 10.1111/imr.12031.

Roberts CW, Walker W, Alexander J. 2001. Sex-associated hormones and immunity to protozoan parasites. *Clinical Microbiology Reviews* **14**:476–488 DOI 10.1128/CMR.14.3.476-488.2001.

Sack LM, Davoli T, Li MZ, Li Y, Xu Q, Naxerova K, Wooten EC, Bernardi RJ, Martin TD, Chen T, Leng Y, Liang AC, Scorsone KA, Westbrook TF, Wong KK, Elledge SJ. 2018. Profound tissue specificity in proliferation control underlies cancer drivers and aneuploidy patterns. *Cell* **173**:499–514 DOI 10.1016/j.cell.2018.02.037.

Sopori M. 2002. Effects of cigarette smoke on the immune system. *Nature Reviews Immunology* **2**:372–377 DOI 10.1038/nri803.

Stampfli MR, Anderson GP. 2009. How cigarette smoke skews immune responses to promote infection, lung disease and cancer. *Nature Reviews Immunology* **9**:377–384 DOI 10.1038/nri2530.

Swann JB, Coquet JM, Smyth MJ, Godfrey DI. 2007. CD1-restricted T cells and tumor immunity. *Current Topics in Microbiology and Immunology* **314**:293–323 DOI 10.1007/978-3-540-69511-0_12.

Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, Sivakumar IKA, Gentles AJ, Malhotra R, Farshidfar F, Colaprico A, Parker JS, Mose LE, Vo NS, Liu J, Liu Y, Rader J, Dhankani V, Reynolds SM, Bowlby R, Califano A, Cherniack AD,

Anastassiou D, Bedognetti D, Mokrab Y, Newman AM, Rao A, Chen K, Krasnitz A, Hu H, Malta TM, Noushmehr H, Pedamallu CS, Bullman S, Ojesina AI, Lamb A, Zhou W, Shen H, Choueiri TK, Weinstein JN, Guinney J, Saltz J, Holt RA, Rabkin CS, Cancer Genome Atlas Research N, Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG, Shmulevich I. 2018. The immune landscape of cancer. *Immunity* **48**:812–830 DOI 10.1016/j.immuni.2018.03.023.

Wang L, Qu J, Liang Y, Zhao D, Rehman FU, Qin K, Zhang X. 2020. Identification and validation of key genes with prognostic value in non-small-cell lung cancer via integrated bioinformatics analysis. *Thoracic Cancer* **11**:851–866 DOI 10.1111/1759-7714.13298.

Wiemels JL, Walsh KM, De Smith AJ, Metayer C, Gonseth S, Hansen HM, Francis SS, Ojha J, Smirnov I, Barcellos L, Xiao X, Morimoto L, McKean-Cowdin R, Wang R, Yu H, Hoh J, DeWan AT, Ma X. 2018. GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21. *Nature Communications* **9**:Article 286 DOI 10.1038/s41467-017-02596-9.

Woodruff PG, Koth LL, Yang YH, Rodriguez MW, Favoreto S, Dolganov GM, Paquet AC, Erle DJ. 2005. A distinctive alveolar macrophage activation state induced by cigarette smoking. *American Journal of Respiratory and Critical Care Medicine* **172**:1383–1392 DOI 10.1164/rccm.200505-686OC.

Xu F, Lin H, He P, He L, Chen J, Lin L, Chen Y. 2020. A TP53-associated gene signature for prediction of prognosis and therapeutic responses in lung squamous cell carcinoma. *Oncoimmunology* **9**:1731943 DOI 10.1080/2162402X.2020.1731943.

Zhang Y, Ou DH, Zhuang DW, Zheng ZF, Lin ME. 2020. In silico analysis of the immune microenvironment in bladder cancer. *BMC Cancer* **20**:265 DOI 10.1186/s12885-020-06740-5.

Zhu Y, Zhang X. 2019. Investigating the significance of tumor-infiltrating immune cells for the prognosis of lung squamous cell carcinoma. *PeerJ* **7**:e7918 DOI 10.7717/peerj.7918.