Prognostic value of immune-related genes and comparative analysis of immune cell infiltration in lung adenocarcinoma of different genders

Tao Fan
National Cancer Center

Chunxiang Li
National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College

Jie He (✉️ prof.jiehe@gmail.com)
National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College

Research

Keywords: Lung adenocarcinoma, Gender, Tumor microenvironment, ESTIMATE algorithm, Tumor-infiltrating immune cells

DOI: https://doi.org/10.21203/rs.3.rs-104556/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background

Lung adenocarcinoma (LUAD) is one of the most important subtypes of lung cancer. Compared with male LUAD patients, female patients have a higher incidence, but better long-term survival rate, with unknown reasons. In this study, we aimed to explore the effect of gender on immune cell infiltration of LUAD tumor microenvironment (TME), and tried to clarify the reasons for the different clinical characteristics of male and female LUAD patients, by conducting a comparative analysis of the TME in female and male LUAD patients.

Methods

We calculated immune and stromal scores of LUAD patients samples from The Cancer Genome Atlas (TCGA) database. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were conducted to reveal biological processes of intersecting genes. Cox regression analysis and protein-protein interaction (PPI) network analysis were conducted to screen immune-related prognostic genes in female (CCR2, LCP2, and PTPRC) and male (BTK and CCR2) patients. Kaplan-Meier survival analysis was used to evaluate prognostic value of identified genes. Mann-Whitney test was used to compare various indicators of male patients and female patients. The main result was subsequently validated in 420 cases from GSE72094.

Results

304 and 368 intersecting genes were identified in female and male patients respectively. The immune score is ranged from −943.17 to 3229.35 among female patients with LUAD and from −541.75 to 3441.78 among male patients with LUAD. The stromal score is ranged from −1790.23 to 2097.27 among female patients and from −1786.94 to 1722.70 among male patients. The immune and stromal scores of women were higher than those of men (P < 0.05). CCR2, LCP2 and PTPRC were identified as the most important immune-related prognostic genes in women. BTK and CCR2 were identified as the most important immune-related prognostic genes in men. Female patients had a higher proportion of memory B cells than that of male patients, while the percentage of T cells CD4 naïve and resting NK cells was lower in female patients (P < 0.05).

Conclusions

This study compared the differences in TIICs in the TME of male and female LUAD patients, and identified prognosis-related genes for patients of different genders.
Background

As the most commonly diagnosed cancer, lung cancer is the leading cause of cancer death with a 5-year overall survival rate of 7.5% in males and 8.5% in females[1, 2]. Among of all histologic subtypes of lung cancer, lung adenocarcinoma (LUAD) is the most prevalent, accounting for more than 40% of lung cancer incidence[3]. There are gender differences in the occurrence and progression of LUAD, but so far the reason is unclear. This study is the first to compare the differences in tumor-infiltrating immune cells (TIICs) in TME between female LUAD patients and male LUAD patients, to explore the influence of gender factors on LUAD.

Recent studies have shown that components in TME, including immune cells, endothelial cells, mesenchymal cells, inflammatory mediators and extracellular matrix, are closely related to the occurrence, progression and prognosis of cancer[4–9]. As the two main components, TIICs and stromal cells are regarded as valuable for the effect of tumor immunotherapy[10–14]. It is reported that immune cells, such as T and B cells, macrophages, dendritic cells, monocytes, etc., interact with tumor cells and are correlated to tumor cell proliferation, metastasis and prognosis[15–17]. Regulated by Tregs and M2 macrophages, tumor-infiltrating lymphocytes (TILs) participate in the clinical course of various cancers[18–22]. Donnem and his colleagues found that the density of CD8 (+) TIL is an independent prognostic factor in non-small cell lung cancer (NSCLC)[23]. In the study of colorectal cancer, it was found that the level of DNA microsatellite instability is related to the density of TIL. High-level of microsatellite instability is characterised by the presence of TILs and a favourable prognosis[24]. Furthermore, TME also has an impact on tumor-related gene expression and clinical prognosis[25–28]. These findings indicate that TME is closely related to tumor progression, and this will provide a potential cure for cancer.

Estimation of STromal and Immune cells in MAignant Tumours using Expression data (ESTIMATE) is created by Yoshihara and his colleagues, which has been used to score the microenvironment of various types of cancer[29]. For example, the ESTIMATE algorithm has been applied to evaluate immune and stromal scores to predict the proportion of immune cells, stromal cells and other non-tumor cells in TME of lung cancer[30], breast cancer[31], uveal melanoma[32], gastric cancer[33], colorectal cancer[34].

As a clinical variable, gender plays a significant role in tumor progression, which should be considered as an important factor in pathogenesis of cancer. Women tend to have lower mortality lower rates of lung cancer than men. In our study, we used gene expression profile data from The Cancer Genome Atlas (TCGA) database to comparatively analyze the difference of TIICs and tumor-related genes between female LUAD patients and male LUAD patients based on ESTIMATE algorithm, in order to explore the causes of higher mortality of male patients with LUAD than female patients.

Methods And Materials

Data Mining from the public Database
Gene expression data of female (n=304) and male (n=247) patients with LUAD were downloaded from the data portal of TCGA (https://tcga-data.nci.nih.gov/tcga/), which included complete clinical information (age, TNM stage, survival and outcome). Estimate score, immune score and stromal score were calculated by ESTIMATE algorithm. 232 female and 188 male LUAD patients data from GSE72094 was downloaded using Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo) to serve as validation sets.

**Identification of differentially expressed genes**

Differentially expressed genes (DEGs) in high- and low- score groups were screened by using the “limma” package [35]. A fold change of >1 and an adjusted p<0.05 was set as the cutoff criteria. Additionally, “heatmap” package was used to construct heat maps.

**Survival analysis**

The correlation of estimate, immune and stromal scores with overall survival of LUAD patients was plotted by “survival” and “survminer” packages.

**Functional analysis and Cox analysis of DEGs**

Functional analysis of DEGs in female and male LUAD patients was conducted using “clusterProfiler”, “org.Hs.eg.db”, “enrichplot” and “ggplot2” packages for GO terms and KEGG pathways. The p value < 0.05 and q value < 0.05 was considered to be statistically significant. At the same time, we performed univariate regression analysis on DEGs to screen prognostic factors that regulate TME and tumor immunity.

**PPI Analysis**

The protein-protein interaction (PPI) networks was analyzed by an online tool named Search Tool for the Retrieval of Interacting Genes (STRING)[36]. Confidence Score≥0.9 was set as the cutoff value. We analyzed the connectivity degree of each DEGs using STRING database and reconstructed the networks via Cytoscape software.

**TME analysis**

CIBERSORT is a deconvolution algorithm that has been widely used to calculate the numbers of each type of TIICs and analyze the correlation between gene expression and TIICs in TME[37-40]. The “corrplot” package was used to conduct correlation-based heatmaps.

**Statistical analysis**

Kaplan-Meier plots were used to analyze and visualize the associations of estimate, immune, and stromal scores and DEGs with prognosis. The correlations of estimate, immune, and stromal scores with age and TNM stage were estimated via Wilcoxon test or Mann-Whitney U tests. The estimate, immune, and
stromal scores of female LUAD patients were compared with that of male LUAD patients using GraphPad Prism 8. Functional analysis, Cox analysis, survival, and TME analysis were conducted using R version 3.6.2. P < 0.05 was considered to be statistically significant.

Results

Comparison of tumor microenvironment scores and gene expression profiles in patients with lung adenocarcinoma (LUAD) of different genders

Gene expression data and corresponding clinical information were downloaded from TCGA database, including 304 cases of female LUAD patients and 247 cases of male LUAD patients. The estimate score is ranged from -2358.46 to 4889.83 among female patients with LUAD and from -2328.69 to 4818.63 among male patients with LUAD. The immune score ranged from -943.17 to 3229.35 among female patients with LUAD and from -541.75 to 3441.78 among male patients with LUAD. The stromal score is ranged from -1790.23 to 2097.27 among female patients with LUAD and from -1786.94 to 1722.70 among male patients with LUAD. The estimate, immune and stromal scores of women were higher than those of male (p<0.05) (Figure 1A). The samples were divided into high-and low-score groups by the medium value. Up-regulated genes and down-regulated genes were showed in Figure 1B. By interaction of DEGs in immune and stromal groups, 269 up-regulated genes and 35 down-regulated genes were screened in female patients, and 340 up-regulated genes and 28 down-regulated genes were screened in male patients (Figure 1C).

Associations of the scores with age, TNM stage and prognosis in LUAD patients of different genders

To explore the relationship between estimate, immune or stromal scores and age, TNM stage or survival, the samples were divided into high-score and low-score groups. The analysis indicated that for female patients with LUAD, T1 stage had higher estimate scores than that of T3 stage (p=0.01), and N1 stage had higher estimate score than that of N0 stage (p=0.015) and N2 stage (p=0.018) (Figure 2A). However, in male patients with LUAD, estimate score had no obvious relationship with the T or N stage (Figure 2D). Whether it was a male or female patient with LUAD, high immune score was related to older age (Figure 2B, 2E). Immune score of T1 stage was higher than that of T3 stage in female patients (Figure 2B), but in male patients, immune score of T1 stage was higher than that of T2 and T4 stage (Figure 2E). In addition, N1 stage had higher immune scores than that of N0 stage (p=0.038) and N2 stage (p=0.011) (Figure 2B) for female patients, and clinical stage I had higher immune scores than that of stage III for male patients (p=0.037) (Figure 2E). For female patients with LUAD, the stromal scores of N1 stage were higher than that of N0 (p=0.016) (Figure 2C), and stromal scores of M0 stage or clinical stage I were higher than that of M1 stage (p=0.044) or clinical stage IV (p=0.024) (Figure 2F). To study the correlations of overall survival with estimate, immune or stromal scores, we performed Kaplan-Meier survival analysis (Figure 2G). The results indicated that high estimate score had better overall survival for female patients with LUAD (p=0.039), but they had no significant relationship with prognosis of male
patients (p=0.039). The overall survival was not associated with immune or stromal score whether it was a male patient or a female patient (p>0.05).

**GO and KEGG analysis of immune-related DEGs**

To further explore the functions of immune-related DEGs, we performed GO and KEGG analysis on these genes. The top 5 enriched BPs in female patients with LUAD based on GO analysis were leukocyte proliferation, lymphocyte proliferation, mononuclear cell proliferation, T-cell activation, and lymphocyte differentiation (Figure 3A), and the top 5 enriched BPs in male patients with LUAD based on GO analysis were T-cell activation, leukocyte migration, leukocyte proliferation, lymphocyte proliferation, and regulation of immune effector process (Figure 3E). Circular plot demonstrated the functional interactions between the BP, CC or MF pathways and genes as extracted from GO (Figure 3B, 3F). Moreover, the top 5 enriched BPs in female patients with LUAD based on KEGG were viral protein interaction with cytokine and cytokine receptor, cytokine-cytokine receptor interaction, chemokine signaling pathway, hematopoietic cell lineage, and staphylococcus aureus infection (Figure 3C), and the top 5 enriched BPs in male patients with LUAD based on KEGG were cytokine-cytokine receptor interaction, viral protein interaction with cytokine and cytokine receptor, hematopoietic cell lineage, chemokine signaling pathway, and graft-versus-host disease (Figure 3G). Circular plot demonstrated the functional interactions between the BP, CC or MF pathways and genes as extracted from KEGG (Figure 3D, 3H).

**PPI network and COX regression analysis of immune-related DEGs**

The STRING tool was used to plot PPT networks of immune-related DEGs, which were regenerated by Cytoscape (Figure 4A, 4B). The top 30 genes with most number of adjacent nodes in female and male patients with LUAD were showed in Figure 4C and 4D. We performed univariate Cox regression analysis to evaluate the prognostic value of immune-related DEGs in female (Figure 4E) and male (Figure 4F) patients with LUAD (genes with p<0.05 displayed in forest plot). Two-way venn diagram comparing the key genes in female (Figure 4G) and male (Figure 4H) cohorts. CCR2, LCP2, and PTPRC were selected as key prognostic factors of female patients with LUAD. BTK and CCR2 were selected as key prognostic factors of male patients with LUAD.

**The expression level of these key TIICs-related genes and their prognostic value in LUAD patients**

We further revealed the expression level of these key TIICs-related genes and their prognostic value in LUAD patients. According to the result, the expression level of CCR2 in tumor tissues of female LUAD patients was not significantly different from that in normal tissue (p>0.05) (Figure 5A and 5B), but female LUAD patient with high level of CCR2 exhibited a better overall survival (p=0.001) (Figure 5C). The expression level of LCP2 in tumor tissues of female LUAD patients was lower than that in normal tissues (p<0.05) (Figure 5D), yet there was no difference in the expression of LCP2 in paired tumors and adjacent normal tissues (p>0.05) (Figure 5E). Interestingly, female LUAD patient with a high level of LCP2 exhibited a better overall survival (p=0.033) (Figure 5F). The expression level of PTPRC in tumor tissues of female LUAD patients was lower than that in normal tissues (p<0.05) (Figure 5G and 5H), and female LUAD
patient with high level of PTPRC exhibited a better overall survival (p=0.006) (Figure 5I). The expression level of BTK in tumor tissues of male LUAD patients was lower than that in normal tissues (p<0.05) (Figure 5J and 5K), and male LUAD patient with high level of BTK exhibited a better overall survival (p=0.035) (Figure 5L). The expression level of CCR2 in tumor tissues of male LUAD patients was lower than that in normal tissues (p=0.035) (Figure 5M), yet there was no difference in the expression of CCR2 in paired tumors and adjacent normal tissues (p>0.05) (Figure 5N). Interestingly, male LUAD patient with a high level of CCR2 exhibited a better overall survival (p=0.011) (Figure 5O). Associations of these key genes expression with age and TNM stage in LUAD patients of different genders were showed in Figure S1.

CIBERSORT for estimating TIICs components in female and male LUAD microenvironment

We further selected 22 types of immune cell for research on immune landscape of TME in LUAD. The selected major immune cell types included B cells, T cells, NK cells, neutrophils, dendritic cells, mast cells, macrophages, et al., which participated in the regulation process of innate immunity and adaptive immunity. Based on CIBERSORT algorithm, we first analyzed the distribution characteristics of the 22 types of immune cells in each LUAD sample (i.e., 270 female patients and 227 male patients) (Figure 6A and 6B). Additionally, the correlation between various TIICs in female and male LUAD patients varied from weak to moderate (Figure 6C and 6D). The correlation matrix indicated that CD8 T cells and M0 macrophages had a strong negative correlation with memory resting CD4 T cells (Cor=-0.47) and resting mast cells (Cor=-0.41) in female LUAD patients. CD8 T cells, M2 macrophages, and resting dendritic cells had a strong positive correlation with memory activated CD4 T cells (Cor=0.49), monocytes (Cor=0.42), and resting mast cells (Cor=0.4) in female LUAD patients (Figure 6C). As shown in Figure 6D for male LUAD patients, there was a significant negative correlation between CD8 T cells and memory resting CD4 T cells (Cor=-0.41), and also a significant positive correlation between CD8 T cells and memory activated CD4 T cells (Cor=0.47).

Difference analysis of TIICs in LUAD tumor and adjacent normal tissues

To compare the differences in immune cell infiltration between female and male patients with LUAD, we analyzed the differences in the distribution of TIICs in the TME of female and male patients with LUAD (Figure 7). Compared with paracancerous tissues in female patients, 8 TIICs (memory B cells, plasma cells, memory activated CD4 T cells, follicular helper T cells, regulatory T cells, gamma delta T cells, M1 macrophages, and resting dendritic cells) had a higher proportion in cancerous tissues, and 6 TIICs (memory resting CD4 T cells, resting NK cells, monocytes, M2 macrophages resting mast cells, eosinophils, and neutrophils) had a lower proportion in cancerous tissues (p<0.05) (Figure 7A). Similarly in male LUAD patients, 5 TIICs (plasma cells, memory activated CD4 T cells, follicular helper T cells, regulatory T cells, and M1 macrophages) accounted for a higher proportion in cancerous tissues than that in paracancerous tissues, and the proportion of 4 TIICs (resting NK cells, monocytes, resting mast cells, and neutrophils) in cancerous tissues was lower than that in paracancerous tissues (p<0.05) (Figure
Female patients with LUAD had a higher proportion of memory B cells, while the percentage of T cells CD4 naïve and resting NK cells was lower in female patients (Figure 7C).

In order to further research the influence of the proportions of TIICs on the prognosis of female and male LUAD patients, univariate Cox regression analysis was conducted on the 22 TIICs for female and male patients respectively. The proportion of activated dendritic cells was identified as risk factor for female LUAD patients (Table S1). The proportion of gamma delta T cells, activated NK cells and activated mast cells were identified as risk factors for male LUAD patients (Table S2).

**Impact of the key identified genes on TIICs**

We further analyzed the effects of the main differentially expressed genes in female (Figure S2) and male (Figure S3) patients with LUAD on immune cell infiltration. For female LUAD patients, the relative proportions of memory B cells (p = 0.004), CD4 memory resting T cells (p = 0.014), resting dendritic cells (p = 0.049) and resting mast cells (p = 0.011) were significantly upregulated in high CCR2 expression group compared with the low CCR2 expression group (Figure S2A). M0 macrophages (p = 0.019) and activated mast cells (p = 0.006) were present in lower proportions in the high CCR2 expression group than in the others (Figure S2A). The relative proportions of CD 8 T cells (p = 0.019), CD4 memory activated T cells (p = 0.005), M1 macrophages (p = 0.009) and neutrophils (p = 0.027) were significantly upregulated in high LCP2 expression group compared with the low LCP2 expression group (Figure S2B). Plasma cells (p = 0.002) and M0 macrophages (p = 0.038) were present in lower proportions in the high LCP2 expression group than in the others (Figure S2B). The relative proportions of memory B cells (p = 0.035), CD4 memory activated T cells (p = 0.013), and eosinophils (p = 0.016) were significantly upregulated in high PTPRC expression group compared with the low PTPRC expression group (Figure S2C). Activated NK cells (p = 0.044) and activated mast cells (p = 0.008) were present in lower proportions in the high PTPRC expression group than in the others (Figure S2C).

For male LUAD patients, the relative proportions of memory B cells (p = 0.022), CD8 T cells (p = 0.034), CD4 memory activated T cells (p < 0.001), monocytes (p < 0.001), resting dendritic cells (p = 0.002), and eosinophils (p < 0.001) were significantly upregulated in high BTK expression group compared with the low BTK expression group (Figure S3A). Plasma cells (p = 0.024), follicular helper T cells (p = 0.031), T cells regulatory (Tregs) (p = 0.039), activated NK cells (p = 0.013), and M0 macrophages (p < 0.001) were present in lower proportions in the high BTK expression group than in the others (Figure S3A). The relative proportions of CD8 T cells (p = 0.008), CD4 memory activated T cells (p < 0.001), monocytes (p = 0.006), and M1 macrophages (p < 0.001) were significantly upregulated in high CCR2 expression group compared with the low CCR2 expression group (Figure S3B). Gamma delta T cells (p = 0.031) and M0 macrophages (p < 0.001) were present in lower proportions in the high CCR2 expression group than in the others (Figure S3B).

**Validation of TCGA results with GEO database**
In order to verify the prognostic value of the identified genes from TCGA, we used GSE72094 as a validation cohort. Patients were divided into high expression groups and low expression groups respectively according to gene expression. We first compared the survival curves of female LUAD patients and male LUAD patients, which showed that the overall survival rate of female patients is better than that of male patients (Figure 8A). Kaplan-Meier survival curves further confirmed that male LUAD patients with high expression levels of CCR2 (Figure 8B) and BTK (Figure 8C) were present in a better overall survival than those with low expression levels of CCR2 and BTK. Similarly, Kaplan-Meier survival curves further verified that female patients with high expression levels of CCR2 (Figure 8D) LCP2 (Figure 8E), and PTPRC (Figure 8F) were present in a better overall survival than those with low expression levels of CCR2 LCP2, and PTPRC.

Discussion

LUAD is the most common subtype of lung cancer. The incidence and mortality of LUAD in female and male patients are different. The prognosis of male LUAD patients is usually worse than female patients, but the cause is currently unknown. Utilizing TCGA database, this study is the first to compare the differences in the infiltration of immune cells in TME between female and male patients with LUAD, which provided in-depth insights for clarifying the reasons for differences in the prognosis of patients of different genders.

The ESTIMATE algorithm has been widely used in recent years to evaluate the immune score and stromal score, which can help scholars understand the TME of LUAD in depth. Our results indicated that the estimate score ranged from -2358.46 to 4889.83 among female patients with LUAD and from -2328.69 to 4818.63 among male patients. Similarly, immune score ranged from -943.17 to 3229.35 among female patients with LUAD and from -541.75 to 3441.78 among male patients. Stromal score ranged from -1790.23 to 2097.27 among female patients with LUAD and from -1786.94 to 1722.70 among male patients. In general, average estimate, immune and stromal scores of female patients with LUAD were higher than those of male patients. Interestingly, except that female patients with high estimate score had better overall survival than that with low estimate score, there was no significant difference in overall survival among other high-score and low-score female or male patients. Early research showed that immune-hot tumors are defined as those in which many immune cells such as T cells, lymphocytes et al. have a high proportion, and these infiltrating immune cells improve efficiency of tumor to response to treatment of immune checkpoint inhibitors[41–43]. More immune-hot tumor means higher immune score in the tumor microenvironment. However, cancer patients with higher immune score do not always mean more immune-hot tumor. In this study, female patients presented higher immune score compared with male patients with LUAD. As showed in Fig. 7, female patients with LUAD had a higher proportion of memory B cells compared with male patients, while the percentage of naïve CD4 + T cells and resting NK cells were lower in female patients, and most immune cells, including naïve B cells, plasma cells, CD8 + T cells, resting/activated memory CD4 + T cells, M0/M1/M2 macrophages, resting/activated dendritic cells, follicular helper T cells, regulatory T cells, gamma delta T cells, activated NK cells, monocytes, resting/activated mast cells, eosinophils, and neutrophils were not significantly different between female
and male patients with LUAD. Hence, although female and male LUAD patients have differences in immune cell infiltration of tumor microenvironment, more study has to be conducted to verify tumor sensitivity to immunotherapy.

In order to further study the difference of immune-related genes between female and male LUAD patients, we divided female (male) LUAD patients into high- and low- score groups. A total of 304 DEGs were identified in female patients and 368 DEGs were identified in male patients (Fig. 1). GO and KEGG analysis for these DEGs indicated that although there were differences in the signaling pathways enriched in female and male patients with LUAD, they were all closely related to tumor immunity (Fig. 3). We further used PPI and Cox methods to screen out the most critical prognostic immune-related genes in female and male LUAD patients. The results showed that CCR2, LCP2, and PTPRC were screened as key prognostic factors of female patients, and BTK and CCR2 were screened as key prognostic factors of male patients with LUAD (Fig. 4).

Next, we compared the infiltration of immune cells in female and male patients with LUAD. Our results indicated 8 TIICs had a higher proportion in cancerous tissues compared with paracancerous tissues in female patients (Fig. 7A), and 5 TIICs accounted for a higher proportion in cancerous tissues than that in paracancerous tissues in male patients (Fig. 7B). Recent studies demonstrate that patients with B cells rich in TME have a better prognosis and immunotherapy response[44–46]. Memory B cells are the basis for humans to have long-lasting immunity. They respond to reencountered antigens by forming germinal centers (GC) and rapidly producing antibodies[47]. It is reported that high density of tumor-infiltrating memory B cells is closely related to superior survival[48–50]. In our study, the density of memory B cells in TME of female patients with LUAD was significantly higher than that in TME of male patients (Fig. 7C), which may explain that female patients with LUAD have a better prognosis than male patients.

Conclusion

We presented a detailed analysis of the tumor microenvironment immune cell infiltration in female and male patients with LUAD. We found differences in the infiltration of immune cells, the expression of prognostic immune-related genes and related signaling pathways in the tumor microenvironment of female and male patients with LUAD. More importantly, female patients with LUAD had a higher proportion of memory B cells compared with male patients, which provides a reliable theoretical basis for explaining the better prognosis of female LUAD patients than that of male patients. Although this study can provide us with insight into the tumor microenvironment of patients with lung adenocarcinoma of different genders, some potential limitations are still to be addressed. For example, the reason that the proportion of memory B cells in the tumor microenvironment of female LUAD patients was significantly higher than that of male LUAD patients was not known, and the study of prognostic immune-related genes lacked biochemical experiments to verify. Further in vivo and in vitro studies on immune cells and genes are urgently needed to clarify the difference between female and male lung adenocarcinoma and provide a theoretical basis for providing effective and personalized tumor immunotherapy.
Abbreviations

LUAD: lung adenocarcinoma

TME: tumor microenvironment

TCGA: The Cancer Genome Atlas

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes

PPI: protein-protein interaction

TIICs: tumor-infiltrating immune cells

TILs: tumor-infiltrating lymphocytes

NSCLC: non-small cell lung cancer

ESTIMATE: Estimation of STromal and Immune cells in MAalignant Tumours using Expression data

GEO: Gene Expression Omnibus

DEGs: differentially expressed genes

STRING: Search Tool for the Retrieval of Interacting Genes

Declarations

Acknowledgements

Not applicable.

Funding

This work was supported by the National Key R&D Program of China (2018YFC1312100), the National Natural Science Foundation of China (81972196), the CAMS Innovation Fund for Medical Sciences (CIFMS) (2017-I2M-1-005, 2019-I2M-2-002), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2018PT32033), the Innovation team development project of Ministry of Education (IRT_17R10).

Authors' contributions

All authors approved this manuscript. TF performed the data analysis. TF and CL designed this research. TF wrote the original draft. JH and CL reviewed and edited this manuscript. JH and CL funded this
research.

**Availability of data and materials**

All the data were collected from TCGA and GEO database. The data sets used in this study are available on request from the corresponding author.

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no conflict of interest.

**References**

1. Siegel RL, Miller KD, Jemal A: *Cancer statistics, 2020*. *CA: a cancer journal for clinicians* 2020, **70**(1):7-30.
2. Baldwin DR: *Improving outcomes in lung cancer patients*. *The Practitioner* 2011, **255**(1745):19-22, 12.
3. Shi J, Hua X, Zhu B, Ravichandran S, Wang M, Nguyen C, Brodie SA, Palleschi A, Alloisio M, Parisenti G et al: *Somatic Genomics and Clinical Features of Lung Adenocarcinoma: A Retrospective Study*. *PLoS medicine* 2016, **13**(12):e1002162.
4. Wu C, Lin J, Weng Y, Zeng DN, Xu J, Luo S, Xu L, Liu M, Hua Q, Liu CQ et al: *Myeloid signature reveals immune contexture and predicts the prognosis of hepatocellular carcinoma*. *The Journal of clinical investigation* 2020.
5. Challoner BR, von Loga K, Woolston A, Griffiths B, Sivamanoharan N, Semiannikova M, Newey A, Barber LJ, Mansfield D, Hewitt LC et al: *Computational image analysis of T-cell infiltrates in resectable gastric cancer: association with survival and molecular subtypes*. *Journal of the National Cancer Institute* 2020.
6. Kazandjian D, Gong Y, Keegan P, Pazdur R, Blumenthal GM: *Prognostic Value of the Lung Immune Prognostic Index for Patients Treated for Metastatic Non-Small Cell Lung Cancer*. *JAMA oncology* 2019.
7. Roberto Moretto, Salvatore Corallo, Antonino Belfiore, Daniele Rossini, Alessandra Boccaccino, Sara Lonardi, Giovanni Centonze, Federica Morano, Marco Maria Germani, Fotios Loupakis et al: *Prognostic impact of immune-microenvironment in colorectal liver metastases resected after triplets plus a biologic agent: A pooled analysis of five prospective trials*. *Eur J Cancer* 2020, **135**:78-88.
8. Miklikova S, Minarik G, Sedlackova T, Plava J, Cihova M, Jurisova S, Kalavska K, Karaba M, Benca J, Smolkova B et al: Inflammation-Based Scores Increase the Prognostic Value of Circulating Tumor Cells in Primary Breast Cancer. Cancers 2020, 12(5).

9. Xu C, Sui S, Shang Y, Yu Z, Han J, Zhang G, Ntim M, Hu M, Gong P, Chen H et al: The landscape of immune cell infiltration and its clinical implications of pancreatic ductal adenocarcinoma. Journal of advanced research 2020, 24:139-148.

10. Zhang Z, Bao S, Yan C, Hou P, Zhou M, Sun J: Computational principles and practice for decoding immune contexture in the tumor microenvironment. Briefings in bioinformatics 2020.

11. Baxevanis CN, Fortis SP, Perez SA: The balance between breast cancer and the immune system: Challenges for prognosis and clinical benefit from immunotherapies. Seminars in cancer biology 2019.

12. Bonanno L, Pavan A, Dieci MV, Di Liso E, Schiavon M, Comacchio G, Attili I, Pasello G, Calabrese F, Rea F et al: The role of immune microenvironment in small-cell lung cancer: Distribution of PD-L1 expression and prognostic role of FOXP3-positive tumour infiltrating lymphocytes. Eur J Cancer 2018, 101:191-200.

13. Zeng H, Zhou Q, Wang Z, Zhang H, Liu Z, Huang Q, Wang J, Chang Y, Bai Q, Xia Y et al: Stromal LAG-3 cells infiltration defines poor prognosis subtype muscle-invasive bladder cancer with immunoevasive contexture. Journal for immunotherapy of cancer 2020, 8(1).

14. Negi N, Griffin MD: Effects of mesenchymal stromal cells on regulatory T cells: Current understanding and clinical relevance. Stem cells (Dayton, Ohio) 2020, 38(5):596-605.

15. Liu C, Hu C, Li Z, Feng J, Huang J, Yang B, Wen T: Helicobacter pylori Systematic profiling of alternative splicing in -negative gastric cancer and their clinical significance. Cancer cell international 2020, 20:279.

16. Pei X, Chen SW, Long X, Zhu SQ, Qiu BQ, Lin K, Lu F, Xu JJ, Zhang PF, Wu YB: circMET promotes NSCLC cell proliferation, metastasis, and immune evasion by regulating the miR-145-5p/CXCL3 axis. Aging 2020, 12.

17. Ugolini A, Tyurin V, Tyurina Y, Tsyganov E, Dontireddy L, Kagan VE, Gabriilovich DI, Veglia F: Polymorphonuclear myeloid-derived suppressor cells limit antigen cross-presentation by dendritic cells in cancer. JCI insight 2020.

18. Verma R, Hanby AM, Horgan K, Verghese ET, Volpato M, Carter CR, Hughes TA: Levels of different subtypes of tumour-infiltrating lymphocytes correlate with each other, with matched circulating lymphocytes, and with survival in breast cancer. Breast cancer research and treatment 2020.

19. Xu K, Fu Y, Han Y, Xia R, Xu S, Duan S, Zhang Z, Li J: Fewer tumour-specific PD-1CD8 TILs in high-risk "Infiltrating" HPV HNSCC. British journal of cancer 2020.

20. Murata K, Nakatsugawa M, Rahman MA, Nguyen LT, Millar DG, Mulder DT, Sugata K, Saijo H, Matsunaga Y, Kagoya Y et al: Landscape mapping of shared antigenic epitopes and their cognate TCRs of tumor-infiltrating T lymphocytes in melanoma. eLife 2020, 9.
21. Togo M, Yokobori T, Shimizu K, Handa T, Kaira K, Sano T, Tsukagoshi M, Higuchi T, Yokoo S, Shirabe K et al: Diagnostic value of F-FDG-PET to predict the tumour immune status defined by tumoural PD-L1 and CD8 tumour-infiltrating lymphocytes in oral squamous cell carcinoma. *British journal of cancer* 2020, 122(11):1686-1694.

22. Peled M, Onn A, Herbst RS: *Tumor-Infiltrating Lymphocytes—Location for Prognostic Evaluation.* *Clinical cancer research : an official journal of the American Association for Cancer Research* 2019, 25(5):1449-1451.

23. Donnem T, Hald SM, Paulsen EE, Richardsen E, Al-Saad S, Kilvaer TK, Brustugun OT, Helland A, Lund-Iversen M, Poehl M et al: *Stromal CD8+ T-cell Density—A Promising Supplement to TNM Staging in Non-Small Cell Lung Cancer.* *Clinical cancer research : an official journal of the American Association for Cancer Research* 2015, 21(11):2635-2643.

24. Michael-Robinson JM, Pandeya N, Cummings MC, Walsh MD, Young JP, Leggett BA, Purdie DM, Jass JR, Radford-Smith GL: *Fas ligand and tumour counter-attack in colorectal cancer stratified according to microsatellite instability status.* *The Journal of pathology* 2003, 201(1):46-54.

25. Theelen WSME, Krijgsman O, Monkhorst K, Kuilman T, Peters DDGC, Cornelissen S, Ligtenberg MA, Willems SM, Blauuwgeers JLG, van Noesel CJM et al: *Presence of a 34-gene signature is a favorable prognostic marker in squamous non-small cell lung carcinoma.* *Journal of translational medicine* 2020, 18(1):271.

26. Marks DK, Gartrell RD, El Asmar M, Boboila S, Hart T, Lu Y, Pan Q, Yu J, Hibshoosh H, Guo H et al: *Akt Inhibition Is Associated With Favorable Immune Profile Changes Within the Tumor Microenvironment of Hormone Receptor Positive, HER2 Negative Breast Cancer.* *Frontiers in oncology* 2020, 10:968.

27. Zhou Y, Han M, Gao J: *Prognosis and targeting of pre-metastatic niche.* *Journal of controlled release : official journal of the Controlled Release Society* 2020.

28. Janina Henze, Frank Tacke, Olaf Hardt, Frauke Alves, Wa’el Al Rawashdeh: *Enhancing the Efficacy of CAR T Cells in the Tumor Microenvironment of Pancreatic Cancer.* *Cancers* 2020, 12(6).

29. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA et al: *Inferring tumour purity and stromal and immune cell admixture from expression data.* *Nature communications* 2013, 4:2612.

30. Qu Y, Cheng B, Shao N, Jia Y, Song Q, Tan B, Wang J: *Prognostic value of immune-related genes in the tumor microenvironment of lung adenocarcinoma and lung squamous cell carcinoma.* *Aging* 2020, 12(6):4757-4777.

31. Oesper L, Satas G, Raphael BJ: *Quantifying tumor heterogeneity in whole-genome and whole-exome sequencing data.* *Bioinformatics (Oxford, England)* 2014, 30(24):3532-3540.

32. Wang Y, Xu Y, Dai X, Lin X, Shan Y, Ye J: *The prognostic landscape of adaptive immune resistance signatures and infiltrating immune cells in the tumor microenvironment of uveal melanoma.* *Experimental eye research* 2020, 196:108069.

33. Liu Y, Wu J, Huang W, Weng S, Wang B, Chen Y, Wang H: *Development and validation of a hypoxia-immune-based microenvironment gene signature for risk stratification in gastric cancer.* *Journal of*
translational medicine 2020, 18(1):201.

34. Lin K, Huang J, Luo H, Luo C, Zhu X, Bu F, Xiao H, Xiao L, Zhu Z: Development of a prognostic index and screening of potential biomarkers based on immunogenomic landscape analysis of colorectal cancer. Aging 2020, 12(7):5832-5857.

35. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research 2015, 43(7):e47.

36. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP et al: STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic acids research 2015, 43:D447-452.

37. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: Robust enumeration of cell subsets from tissue expression profiles. Nature methods 2015, 12(5):453-457.

38. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D et al: Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature biotechnology 2019, 37(7):773-782.

39. Tamminga M, Hiltermann TJN, Schuuring E, Timens W, Fehrmann RS, Groen HJ: Immune microenvironment composition in non-small cell lung cancer and its association with survival. Clinical & translational immunology 2020, 9(6):e1142.

40. Bense RD, Sotiriou C, Piccart-Gebhart MJ, Haanen JBAG, van Vugt MATM, de Vries EGE, Schröder CP, Fehrmann RSN: Relevance of Tumor-Infiltrating Immune Cell Composition and Functionality for Disease Outcome in Breast Cancer. Journal of the National Cancer Institute 2017, 109(1).

41. Galon J, Bruni D: Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nature reviews Drug discovery 2019, 18(3):197-218.

42. Mlecnik B, Tosolini M, Kirilovsky A, Berger A, Bindea G, Meatchi T, Bruneval P, Trajanoski Z, Fridman WH, Pagès F et al: Histopathologic-based prognostic factors of colorectal cancers are associated with the state of the local immune reaction. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2011, 29(6):610-618.

43. Thomas F. Gajewski, Leticia Corrales, Jason Williams, Brendan Horton, Ayelet Sivan, Stefani Spranger: Cancer Immunotherapy Targets Based on Understanding the T Cell-Inflamed Versus Non-T Cell-Inflamed Tumor Microenvironment. Adv Exp Med Biol 2017, 1036:19-31.

44. Petitprez F, de Reyniès A, Keung EZ, Chen TW, Sun CM, Calderaro J, Jeng YM, Hsiao LP, Lacroix L, Bougoüin A et al: B cells are associated with survival and immunotherapy response in sarcoma. Nature 2020, 577(7791):556-560.

45. Helmink BA, Reddy SM, Gao J, Zhang S, Basar R, Thakur R, Yizhak K, Sade-Feldman M, Blando J, Han G et al: B cells and tertiary lymphoid structures promote immunotherapy response. Nature 2020, 577(7791):549-555.

46. Anne Montfort, Oliver Pearce, Eleni Maniati, Benjamin G. Vincent, Lisa Bixby, Steffen Böhm, Thomas Dowe, Edmund H. Wilkes, Probir Chakravarty, Richard Thompson et al: A Strong B-cell Response Is
Part of the Immune Landscape in Human High-Grade Serous Ovarian Metastases. *Clinical cancer research: an official journal of the American Association for Cancer Research* 2017, 23(1):250-262.

47. Shogo Takatsuka, Hiroyuki Yamada, Kei Haniuda, Hiroshi Saruwatari, Marina Ichihashi, Jean-Christophe Renauld, Daisuke Kitamura: **IL-9 receptor signaling in memory B cells regulates humoral recall responses.** *Nat Immunol* 2018, 19(9):1025-1034.

48. Zhang Z, Ma L, Goswami S, Ma J, Zheng B, Duan M, Liu L, Zhang L, Shi J, Dong L et al: **Landscape of infiltrating B cells and their clinical significance in human hepatocellular carcinoma.** *Oncoimmunology* 2019, 8(4):e1571388.

49. Asmaa M. Zahran, Asmaa S. Shaltout, Hussein Fakhry, Salah M. Khallaf, Ola N. Abdel Fattah, Doaa F. Temerik, Amal Rayan: **Prognostic Significance of Circulating CD28 Negative Suppressor T Cells and Memory B Cells in Patients with Breast Cancer.** *Iran J Immunol* 2020, 17(2).

50. Norouzian M, Mehdipour F, Balouchi Anaraki S, Ashraf MJ, Khademi B, Ghaderi A: **Atypical Memory and Regulatory B Cell Subsets in Tumor Draining Lymph Nodes of Head and Neck Squamous Cell Carcinoma Correlate with Good Prognostic Factors.** *Head and neck pathology* 2019.