



Recognition of immune-related tumor antigens and immune subtypes for mRNA vaccine development in lung adenocarcinoma



Deze Zhao^{a,b,1}, Xianyu Liu^{a,b,1}, Yunhan Shan^{a,b}, JiaJia Li^{a,b}, Weifang Cui^{a,b}, Jun Wang^{a,b}, Junjie Jiang^{a,b}, Qun Xie^c, Chunfang Zhang^{a,b}, Chaojun Duan^{a,b,d,e,*}

^a Department of Thoracic Surgery, Xiangya Hospital, Central South University, Xiangya Road 87th, Changsha 410008, Hunan, PR China

^b Hunan Engineering Research Center for Pulmonary Nodules Precise Diagnosis & Treatment, Changsha 410008, Hunan, PR China

^c Department of Ultrasonic Imaging, Affiliated Hospital of Hunan Traditional Chinese Medicine Research Institute, Changsha 410006, Hunan, PR China

^d National Clinical Research Center for Geriatric Disorders, Changsha 410008, Hunan, PR China

^e Institute of Medical Sciences, Xiangya Lung Cancer Center, Xiangya Hospital, Central South University, Changsha 410008, Hunan, PR China

ARTICLE INFO

Article history:

Received 2 May 2022

Received in revised form 30 August 2022

Accepted 30 August 2022

Available online 5 September 2022

Keywords:

mRNA vaccines

Lung adenocarcinoma

Immune microenvironment

ABSTRACT

Background: There are currently no treatments targeting the immune microenvironment (TME) as an extension of immunotherapy. Our research aims to provide guidance for the development of immune-related mRNA vaccines and the identification of immune subtypes for vaccine treatment in lung adenocarcinoma (LUAD).

Methods: HTRNA-Seq and single cell RNA-seq data were obtained from The Cancer Genome Atlas (TCGA) and Gene-Expression Omnibus (GEO, GSE87340, GSE140343, GSE148071) databases. Immune checkpoints (ICP) were used as criteria to differentiate immune subtypes and immune resistance score (IRS) system is constructed by ssGSEA to judge the immune microenvironment status of patients.

Results: Two overexpressed tumor-specific antigens, including ZC3H12D and TXNDC5, were found to be associated with both disease-free survival (DFS) and overall survival (OS). In addition, the expression of two genes correlated with antigen-presenting cell (APC) infiltration and tumor purity. Subsequently, the immune subtype of the patient was defined by constructing an IRS scoring system. The lower the IRS, the stronger the immune response in the TME. This result was verified in external datasets and at the single-cell level.

Conclusions: ZC3H12D and TXNDC5 are potential tumor-specific antigens for developing mRNA vaccines in LUAD. Importantly, patients with low IRS are more suitable for the use of immunotherapy and vaccines. Our research enhances understanding of TME features and guides more effective immunotherapy strategies.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Worldwide, lung cancer remains the deadliest and the third incidence malignancy, and as global tobacco use increases, especially in Asia, mortality worldwide will continue to rise [1,2]. From 2014 to 2018, mortality of lung cancer, especially non-small cell lung cancer (NSCLC), declined rapidly due to molecular targeted therapies and immunotherapies, and 2-year relative survival increased by 5% to 6% for each stage of diagnosis in NSCLC [1,3]. Although the treatment pattern of NSCLC has been changed

dramatically by immunotherapy, PD-1/L1 monotherapy is strictly selective for patients, with significant benefit in patients with >50% of cells expressing PD-L1 [4,5]. Thus, combination immunotherapy and distinguishing immune subtypes suitable for immunotherapy have become the direction of current oncology research. Currently, the combination therapy is mostly combined with traditional chemotherapy and targeted drugs, and there is no treatment specifically targeting the immune microenvironment as an extension of immunotherapy and no effective biomarkers have been defined to differentiate patients' immune subtypes [6].

Cancer vaccines could induce de novo response to against tumor cells that express tumor specific antigen, mainly including four types (peptides, tumor cells, dendritic cells and genetic DNA/RNA) [7]. Until the 21st century, the application of mRNA

* Corresponding author at: Department of Thoracic Surgery, Xiangya Hospital, Central South University, Xiangya Road 87th, Changsha 410008, Hunan, PR China.

E-mail address: duancjxy@csu.edu.cn (C. Duan).

¹ Contributed equally to this work.

vaccine was limited by the developing mRNA synthesis technology and delivery systems. With these major technological breakthroughs, mRNA have becoming potential treatment strategy of cancers [8]. Since they are not integrated into the genome, mRNA vaccines do not have the potential risk of insertional mutations and can be degraded by RNases in vivo [9]. To date, clinical trials of mRNA vaccine have been conducted against prostate cancer, gastrointestinal cancer, blood cancer, showing a potential efficacy. However, the mRNA vaccine for lung adenocarcinoma (LUAD) is still uncertain due to the tumor heterogeneity, tumor microenvironment (TME) and other factors. Therefore, there is an urgent need to find tumor-specific antigens in LUAD to change the landscape of treatment, either by mono-vaccine or in combination with immunotherapy. At present, vaccine application faces two major challenges, one is identifying specific antigens of tumor cells, and the other is avoiding immune cell depletion caused by unfavorable properties of the TME [10]. The workflow of our study was shown in Fig. S1A. In our study, two tumor-specific antigens were found, and their expression was significantly correlated with immune cell infiltration. Besides, two immune subtypes were defined (immune-resistant and immune-active), which were validated in external datasets and at the single-cell level. Our research provides guidance for the development of immune-related mRNA vaccines and the identification of immune subtypes for vaccine treatment.

2. Materials and Methods

2.1. Data Extraction and Pre-processing

HTRNA-Seq, simple nucleotide variation, copy number variation data and related clinical information of 501 LUAD patients were collected from The Cancer Genome Atlas (TCGA) database.

78 primary LUAD tumor samples and 71 adjacent normal lung tissue samples were obtained from Gene-Expression Omnibus (GEO, GSE87340, GSE140343) as the validation cohort. Furthermore, samples lacking clinical data were excluded, and the FPKM values were translated to $\log_2(\text{FPKM} + 1)$. The Single cell RNA-seq data was downloaded from GEO cohort (GSE148071) and processed by “Seurat” package [11].

“Maftools” package was applied to find the mutant gene [12]. “GISTIC2” software was used to find the gene with copy number amplification and their corresponding chromosome position. Gene set enrichment analysis (GSEA) was performed based on the expression of selected antigens in the R package “GSVA” [13].

2.2. Tumor immune infiltrating analysis

The immune infiltrating cells (especially antigen-presenting cells, APC) and tumor purity was estimated using TIMER and ESTIMATE analysis by “IOBR” package (version 0.99.9) [14]. The association between the immune infiltrating cells and the expression of the selected antigens was investigated through spearman correlation analysis and $P < 0.05$ was considered statistical significance.

The 28 immune signatures estimating infiltrating immune cells were quantified using single-sample gene set enrichment analysis (ssGSEA) by the “GSVA” package [15].

2.3. Human protein Atlas analysis for immune-related antigens

Immunohistochemical staining of selected antigen was download from The Human Protein Atlas (HPA) database and investigate the protein differential expression among normal tissue and LUAD tissue [16].

2.4. Identification of immune subtypes

77 Immune checkpoint genes (ICGs) were collected from a review of the literature [17–20]. Univariate Cox regression was performed based on OS to find prognosis related ICGs. Then, enrichment score (ES) of risk ICGs ($\text{HR} > 1$) and protect ICGs ($\text{HR} < 1$) was calculated using single sample gene set enrichment analysis (ssGSEA) in the R package ‘GSVA’. The immune resistance score (IRS) was defined by the difference in ssGSEA score through risk ICGs minus protect ICGs ES. The IRS was also validated in two individual GEO dataset.

$$\text{IRS} = \text{ssGSEA_ES}(\text{risk_ICGs}) - \text{ssGSEA_ES}(\text{protect_ICGs}).$$

2.5. Drug response analysis

To explore potential drugs for different subtypes of patients based on IRS, we collected drug response data from the Cancer Therapeutics Response Portal (CTRP) and Genomics of Drug Sensitivity in Cancer database (GDSC). Besides, the corresponding data of clinically actionable genes (CAGs) targeted by FDA-approved drugs was obtained from a previous study. The relationship between the drug response and IRS-subtype was investigated through spearman correlation analysis and Wilcoxon test analysis ($P < 0.05$).

2.6. Statistical analysis

“Matched-samples” *t*-test was used to identify over-expressed (OE) genes based on the criterion: the $\log_2\text{FC} > 0.58$ and $\text{adj. } p < 0.05$. Then, univariate Cox regression was performed based on disease-free survival (DFS) and overall survival (OS) to find risk OE-genes and clinical features. Next, prognosis-related clinical features and each prognostic-related gene were combined individually and analyzed by multivariate Cox regression to identify independent risk factors. Statistical analyses were performed using R software (version 4.1.2).

3. Results

3.1. Identification of candidate antigens of LUAD

Tumor-specific gene changes play an important role in tumor development and can be used as potential antigens for vaccine applications. For finding potential antigens of LUAD, the aberrantly overexpression genes were analyzed firstly. As shown in Fig. 1A, 7928 genes were overexpression in TCGA dataset. In addition, another two datasets were considered to identify more credible tumor-specific genes. 1967 overexpressed genes were identified, of which 255 genes were associated with patient prognosis (Fig. 1B). Subsequently, we combined each of the previously identified prognostic genes and prognostic-related clinical features (Fig. 1C) for multivariate regression analysis to identify more stable independent prognostic factors. Ultimately, 88 independent risk genes were considered as tumor-specific antigens for vaccine development.

Next, we analyzed the relationship between these tumor-specific antigens and immune cell infiltration for constructing vaccine that can be used in combination with immunotherapy. Interestingly, 69 of the 88 tumor-specific antigens were associated with at least one antigen-presenting cell infiltration, of which ZC3H12D and TXNDC5 was negatively correlated with tumor purity and positively correlated with infiltration of B cell, Macrophage and Myeloid dendritic cell (Fig. 1D-E). This mean that these tumor-specific

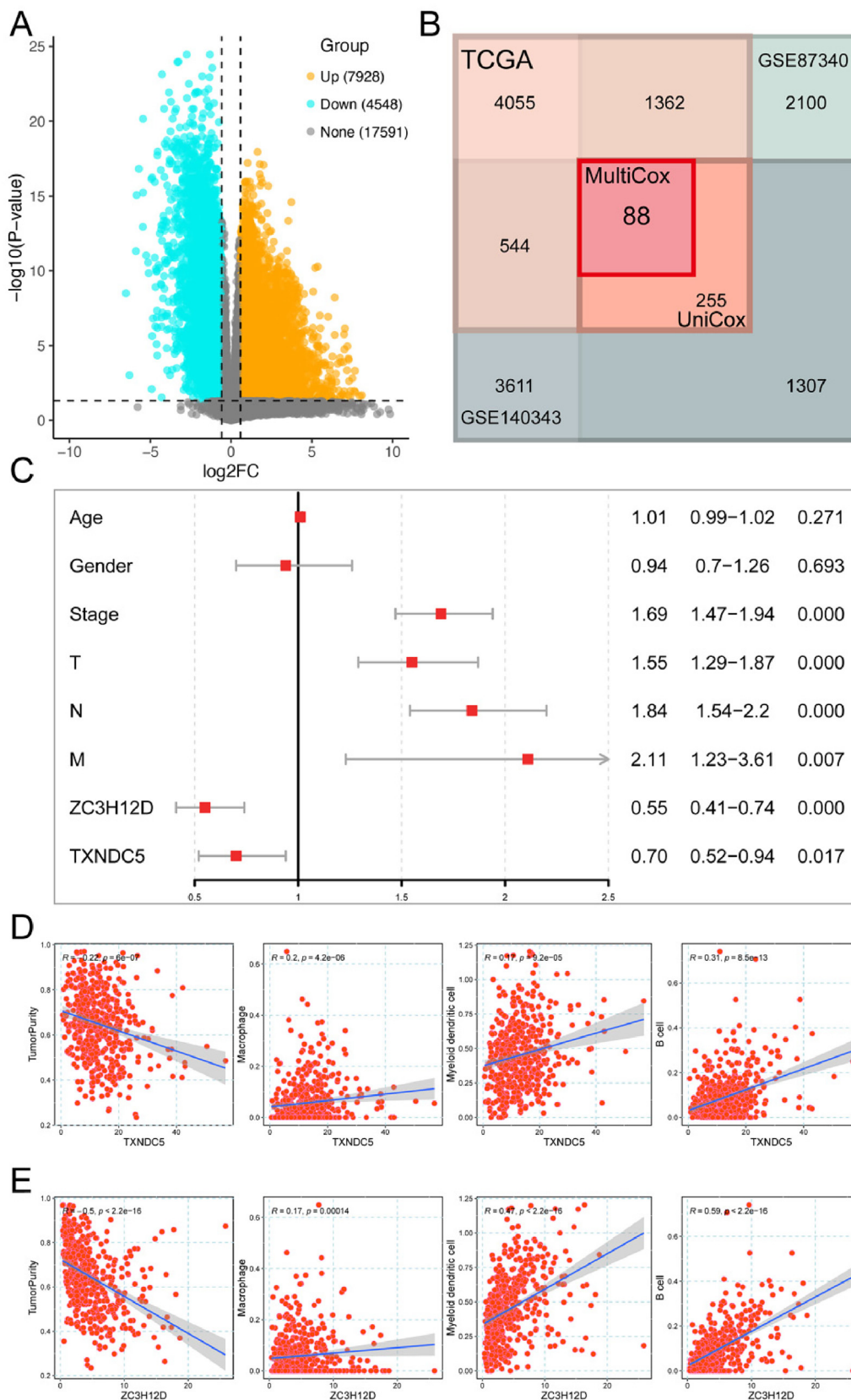


Fig. 1. Identification of potential tumor antigens in lung adenocarcinoma. (A) Volcano plot of gene expression in TCGA cohorts. (B) Venn diagrams for genetic screening: overexpressed and independent risk tumor-specific antigens in three datasets. (C) Forest plot of clinical features. (D-E) Expression correlation between tumor purity, infiltration of B cell, Macrophage and Myeloid dendritic cell and TXNDC5 (D) and ZC3H12D (E).

antigens could be identified by APCs and presented to T cells and captured by B cells to simulate immune response.

Subsequently, we explored the protein level of ZC3H12D and TXNDC5 in LUAD (Fig. 2A-B). Consistent with mRNA expression levels, the immunohistochemistry showed that the protein expression of TXNDC5 in adenocarcinoma tissue was significantly higher than that in normal tissue. Besides, the location of ZC3H12D and TXNDC5 protein expression was in both cytoplasm and cell membrane (Figs. 2B and S3A).

In addition, we explored whether ZC3H12D/TXNDC5 are altered in genome reconstitution that contributes to tumorigenesis. First, we analyzed its genetic mutations in the TCGA mutation data but found no specific mutations occurred. Subsequently, we explored the copy number variation of ZC3H12D/TXNDC5 and found that the increased or decreased copy number variation had little effect on the expression of ZC3H12D (Fig. S2A), while the copy number variation of TXNDC5 significantly affected its expression (Fig. S2B). However, the expression of ZC3H12D/TXNDC5 remained higher in the copy number deletion group of tumors than in normal tissues. This means that gene mutation and copy number variation are not the reasons why two genes become tumor-specific antigens, and the specific reasons still need further exploration.

3.2. The function of tumor-specific antigens

We further explored the relationship between immune-related antigens and patient overall survival and disease-free survival. After being divided into high and low expression groups according to gene expression, patients in the high expression group had better prognosis and longer disease-free survival (Fig. 2C-E). In addition, the expression of ZC3H12D was highly in early-stage patients, while the expression of TXNDC5 was not significantly correlated with the tumor stage of patients (Fig. S3B-E). This suggested that the expression of these immune-related antigens was beneficial to the prognosis of patients.

To explore the biological pathways by which immune-related antigens affect tumor development, we performed GSEA analysis on groups of high- and low-expressing antigens (Fig. 2G-H). The results showed that the high expression of both immune-related antigens could result in a significant enrichment of the immune-related pathways, among which TXNDC5 was mainly related to humoral immunity such as B cell receptor signaling pathway, while ZC3H12D is not only related to humoral immunity, but also to cell-mediated immunity such as T cell receptor signaling pathway.

3.3. Identification potential immune subtypes of LUAD

Due to the heterogeneity of the tumor immune microenvironment, there are large differences in the response of different patients to immunotherapy, such as immune cell depleted patients who are not sensitive to the stimulation of immunotherapy. Therefore, distinguishing the immune subtypes of patients and finding patients who respond to immunotherapy is of great significance for guiding the use of mRNA vaccines. We performed univariate Cox regression analysis on 77 immune checkpoints, of which 12 checkpoints were considered as protective factors and 4 checkpoints were risk factors (Fig. 3A). Because of differences in the roles of checkpoints in influencing tumor development, it was not reasonable to combine all checkpoints for patient subtype grouping in previous studies. So, we used GSEA to calculate the risk score and protection score separately based on the expression of prognosis-related checkpoints. Finally, the IRS is obtained by risk score minus the protection score.

Subsequently, we divided the patients into two immune subtype groups based on the median of IRS. As shown in Fig. 3B, risk

checkpoints are highly expressed in group 2, while protective checkpoints are highly expressed in group 1. By comparing the OS of the immune subtype groups, we found that the prognosis of group 1 was significantly better than group 2 (Fig. 3C). Moreover, the weight of proportion for age, clinic stage and TNM stage in the immune subtype groups were also shown in Fig. 3C, which indicated that patients in group 2 had higher probability of tumor progression and recurrence. The predictive stability of the IRS was also validated in two external datasets, and consistent with previous results, patients with low IRS scores had better prognosis (Fig. S4).

Comparing the differences in tumor hallmark pathways between the immune subtype groups, we found that multiple immune-related pathways were significantly enriched in group 1, while proliferation and metabolic pathways were significantly enriched in group 2 (Fig. 3D).

Tumor immunogenic cell death (ICD) is a kind of tumor cell death which dying tumor cells releases damage-associated molecular patterns and leads to tumor-specific immune responses. We explored the expression of ICD gene in immune subtype groups, and found that many immunogenic biomarkers including ANXA1, CXCL10, TLR, were significantly highly expressed in group 1 (Fig. 4A), which mean that the immune response was more active in group 1 resulting in immunogenic death of tumor cells and improved patient outcomes.

Therefore, group 2 with higher IRS scores can be classified as immune resistance subtype, which is characterized by a lack of immune response and vigorous tumor proliferation; group 1 with a lower IRS score can be classified as immune active subtype, which had a stronger immune response and thus may serve as a suitable subtype for mRNA vaccines and immunotherapy.

3.4. Immune microenvironment features of IRS-related immune subtype

To further explore specific differences in the immune microenvironment between immune subtypes, we compared stromal score, immune scores, estimate score and tumor purity between the two groups. The three immune-related scores of group 1 were significantly higher than those of group B, while the tumor purity was lower than that of group 2 (Fig. 4B). Next, we compared immune cell infiltration between the two groups, and consistent with previous results, most immune cells (especially active B cells and CD4+/CD8+ T cells) were significantly enriched in group 1 (Fig. 4C).

In addition, we explored the expression of ZC3H12D and TXNDC5 in immune subtype groups. Consistent with our hypothesis, expression of both genes was increased in group 1, which may partly explain the increased infiltration of immune cells that stimulated immune responses in group 1 (Fig. 4D-E). Overall, the above results showed that group 1 may have a stronger immune response as a hot tumor, in contrast, group 2 as a cold tumor lacks immune cell infiltration, which in turn may be less responsive to immunotherapy and mRNA vaccines.

3.5. Drug response of IRS-related immune subtype

Moreover, we performed drug sensitivity analysis to explore potential treatments targeting different immune subtypes. Firstly, 20 clinical application genes were found altered between two immune subtype groups, and there are four classes of drugs available for treatment, including targeted therapy, immunotherapy, hormone therapy, and chemotherapy (Fig. 5A). Interestingly, group 2 had few clinically operable genes, including proliferation-related genes such as VEGFA, KRAS, and ERBB3, while group 1 had rela-

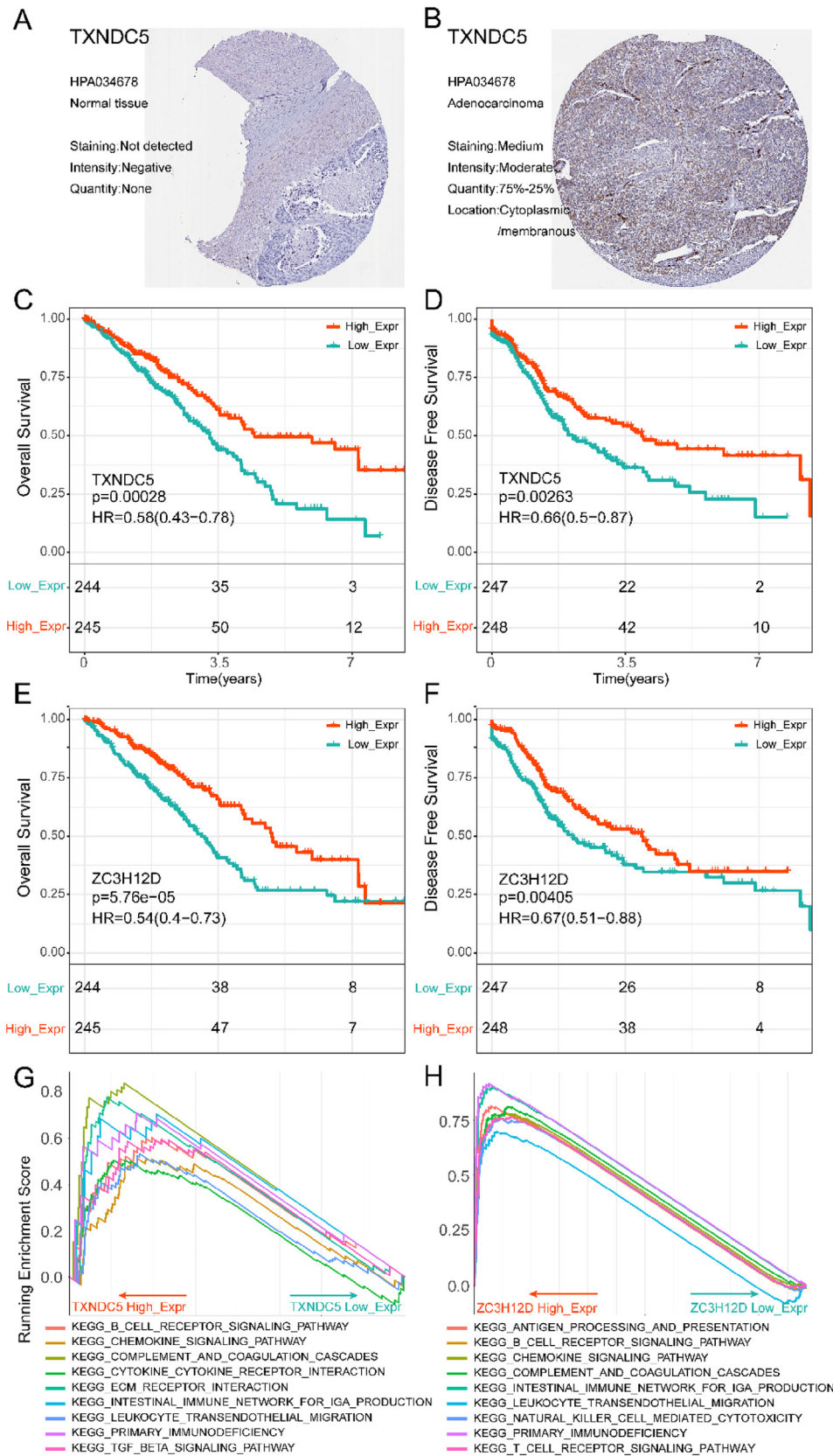


Fig. 2. Characteristics of tumor antigens and potential functional pathways. (A-B) IHC images of TXNDC5 in normal lung tissues (A) and adenocarcinoma tissues (B). (C-F) K-M curves showed the overall survival (OS) and disease-free survival (DFS) of patients with LUAD in the different expression levels of TXNDC5 (C, D) and ZC3H12D (E, F). (G-H) Potential functional pathways associated with TXNDC5 (G) and ZC3H12D (H) expression.

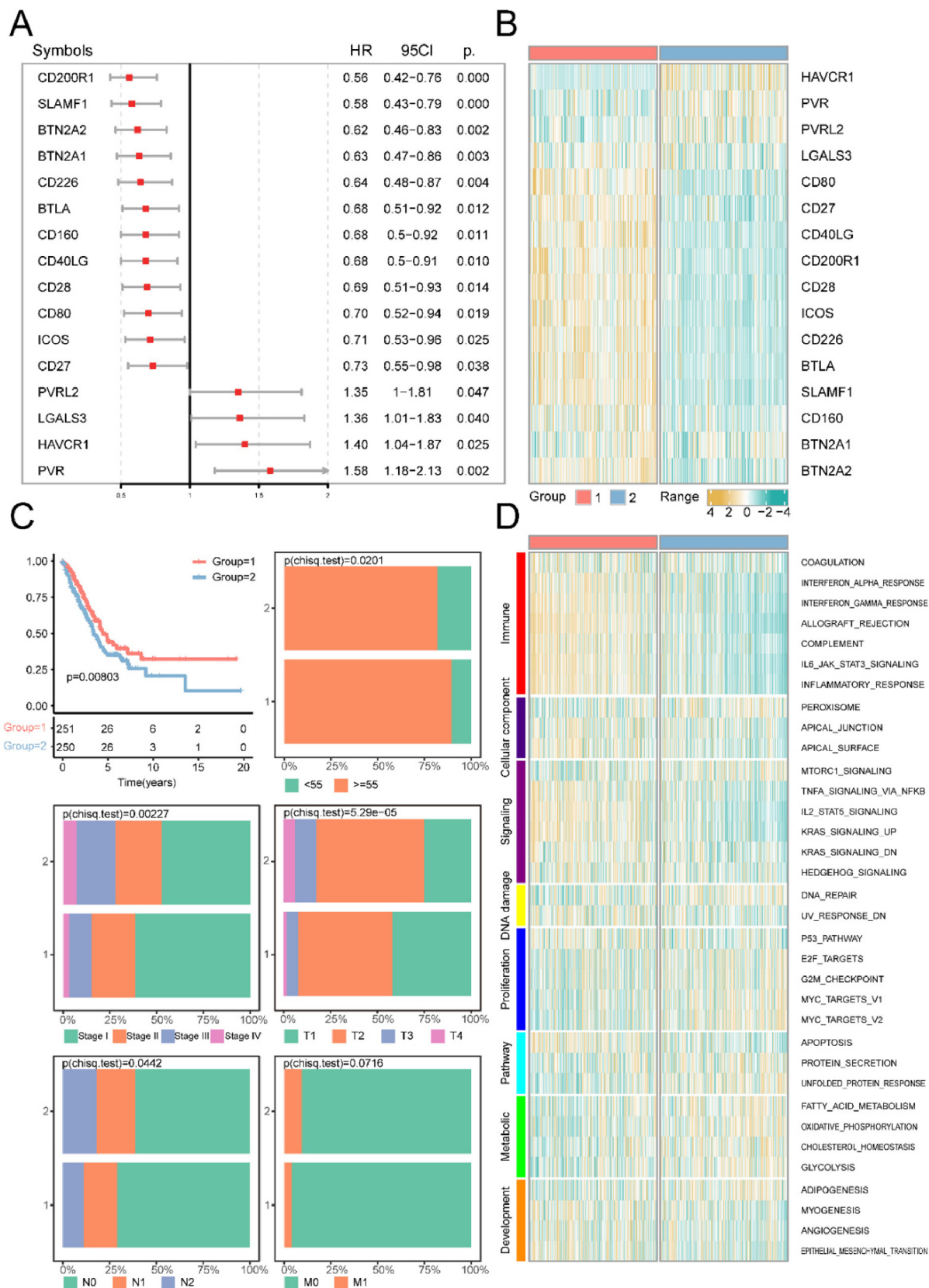


Fig. 3. Identification potential two immune subtypes of LUAD. (A) Forestplot of prognosis-related 16 immune checkpoints, including 4 risk checkpoints and 12 protect checkpoints. (B) Expression of 16 immune checkpoints in two immune subtypes. (C) Differences in clinical features between the two immune subtypes, including prognosis, age, clinic stage and TNM stage. (D) Differences in hallmark pathways between two immune subtypes.

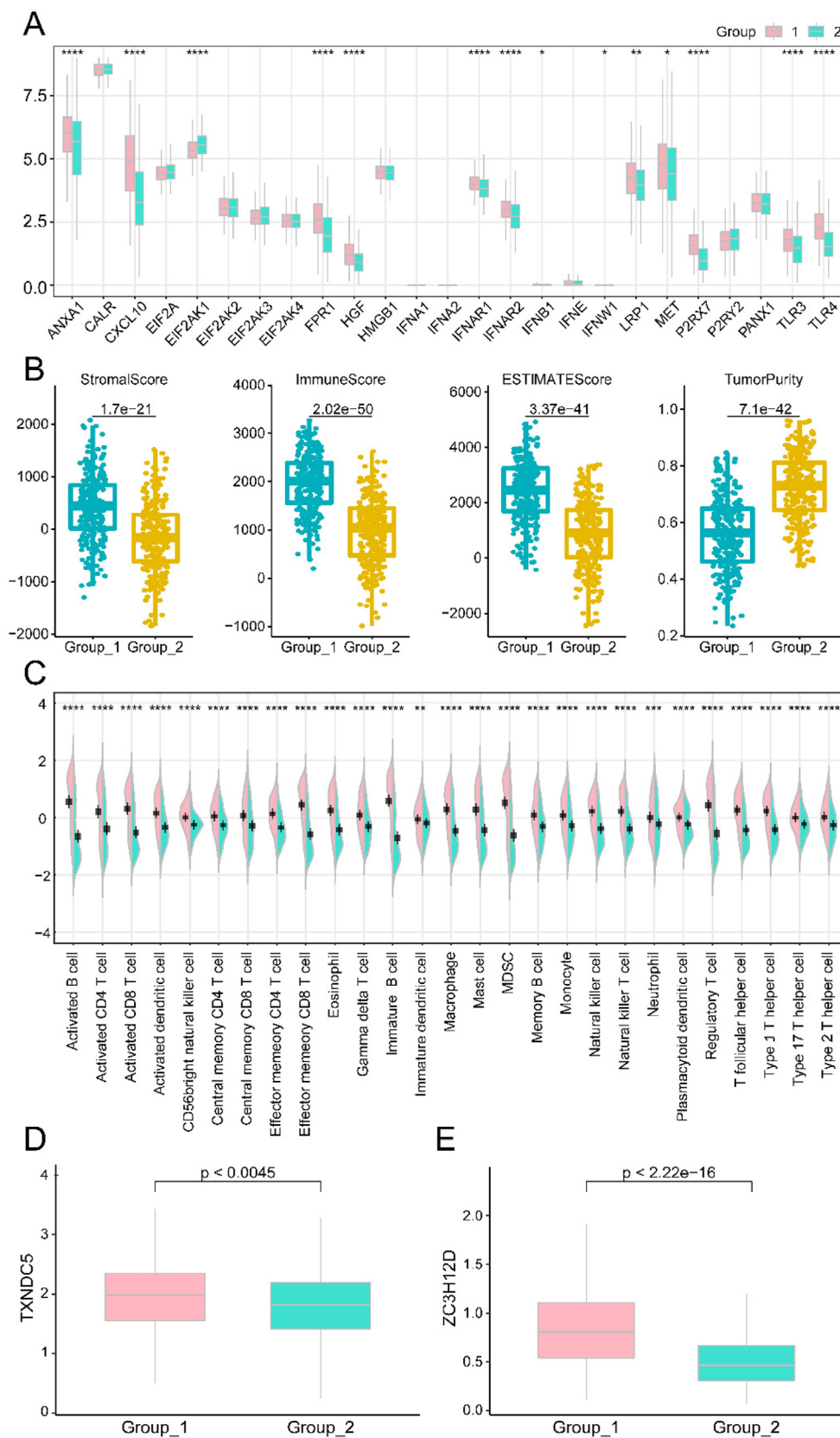


Fig. 4. Characteristics of immunogenic cell death (ICD) genes and tumor immune microenvironment (TIME) in different immune subtypes (A) Difference in the expression of ICD genes between the two immune subtypes. (B) Difference in immune-related scores and tumor purity between the two immune subtypes. (C) Difference in infiltration of immune cell between the two immune subtypes. (D) Differences in expression of TXNDC5 and ZC3H12D between two immune subtypes.

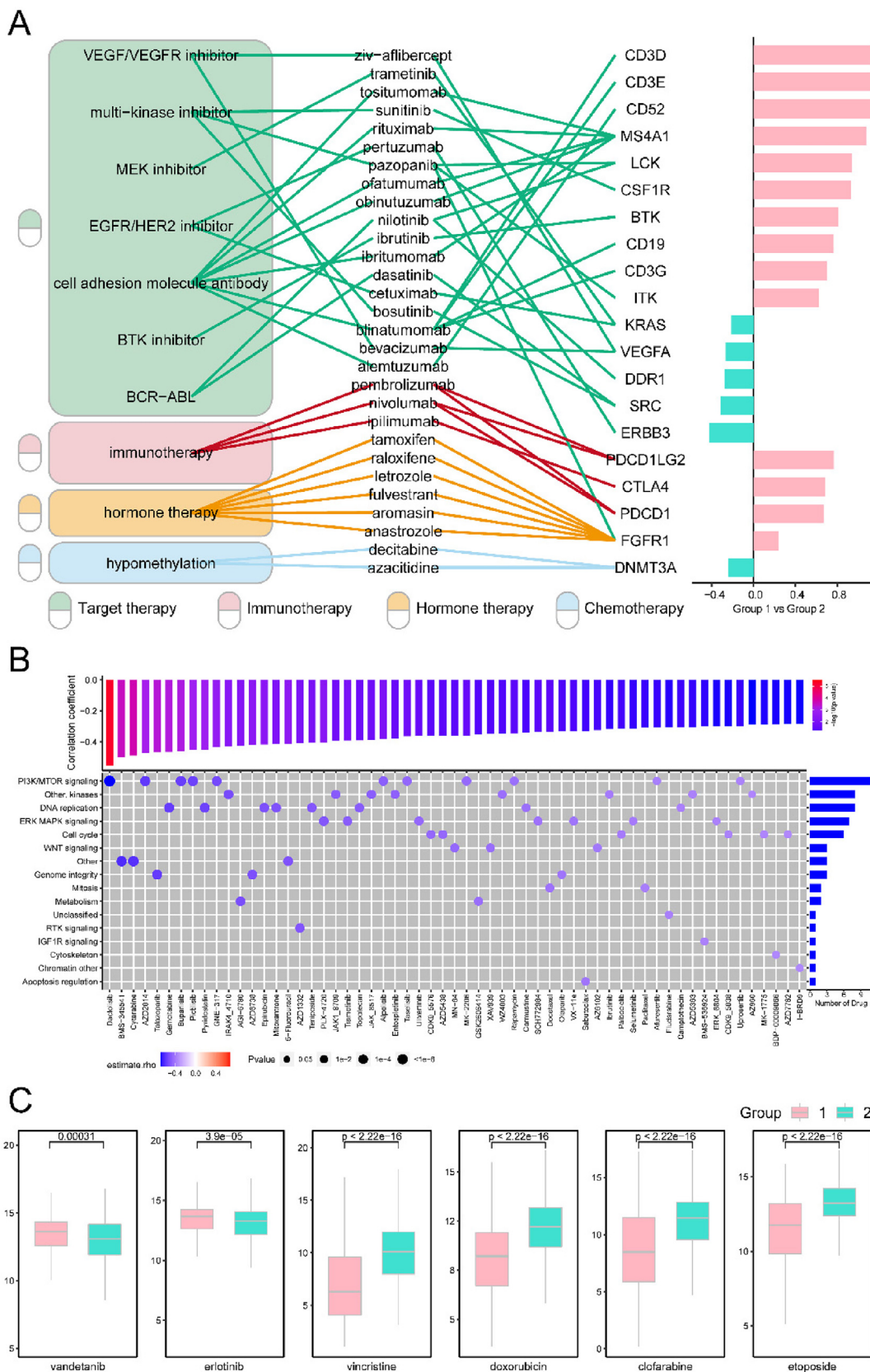


Fig. 5. Associations between two immune subtypes and clinically actionable genes and drug susceptibility. (A) Differentially expressed clinical actionable genes (CAGs) targeted by FDA-approved drugs. The barplot shows CAGs that vary significantly between immune subtype groups. (B) Drugs sensitive to IRS and the signaling pathways they target. The upper barplot represents the correlation between IRS and drug sensitivity. The right histogram represents the number of drugs targeting each pathway. (C) Sensitivity of two immune subtypes to drugs (grouping 835 cancer cell lines from CTRP database based on IRS).

tively more clinically actionable genes, including immune-related genes PDCD1, PDCD1LG2, CTLA4, etc.

To further explore IRS-related actionable drugs, we obtained gene expression and drug response data from GDSC and CTRP. Spearman correlation analysis between the FPS and drug response in cancer cell lines was performed ($|R_s| > 0.25$; $FDR < 0.05$), and 58 drugs were found to be IRS-sensitive, while none were IRS-resistant. Interestingly, most of the 58 drugs were proliferation-related pathways, including PI3K/MTOR signaling pathway, ERK MAPK signaling pathway and DNA replication pathway, which is consistent with previous results that proliferation-related drugs are more suitable for patients with high IRS scores (Fig. 5B). We then divided the CTRP database cells into two groups based on the IRS and explored which FDA drugs were available for the two groups of patients. The results show that vandetanib and erlotinib were suitable for patients with low IRS, while vincristine, doxorubicin, clofarabine and etoposide were suitable for patients with high IRS (Fig. 5C).

In summary, we predicted IRS-related clinical drugs, which can guide the design of rational combination vaccines and drug regimens for patients with different immune subtypes.

3.6. Analysis of IRS-related immune subtype at the single-cell level

Compared with traditional tissue sequencing, single-cell profiling technology could better characterize intra-tumor heterogeneity, so we further explored IRS scoring system at single-cell resolution (Fig. S5). Using graph-based principal component clustering combined with marker-based annotation, we divided cells in GSE148071 into 10 clusters (Fig. 6A).

Subsequently, we calculated the IRS of each cell and found that IRS was generally highest in malignant cells, followed by epithelial cells, while the lowest IRS was observed in immune cells such as T and B cells (Fig. 6B). Then, the IRS level (median IRS of all cells) and the proportion of immune cells were calculated for each patient. As shown in Fig. 6C, the proportion of tumor cells increased with IRS, which means that the IRS can also reflect the patient's immune cell infiltration at the single-cell level, that is, with the increase of IRS, the immune cell reduction. In addition, we also divided patients into immune resistance and active subtype based on IRS. As expected, the difference in cell proportions between the two immune subtypes was statistically significant. Anti-tumor immune cells such as NK cells, B cells, effector CD8⁺ T cells, and CXCL13⁺ CD4 T cells were lower in the immune-resistant subtype, indicating that TME may be resistant to immune responses and thus can be classified as cold tumors (Fig. 6D). Due to the limited sequencing depth, we did not detect ZC3H12D and TXNDC5 expression in single cells and therefore could not compare its expression in immune subtypes.

Further, we analyzed the cellular interactions within the immune microenvironment of the two immune subtypes. Overall, the immune active subtypes have abundant cellular interactions in TME. On the contrary, the immune resistance subtypes obviously lack cell interactions, and most interactions are immunosuppressive, such as M2 macrophage, and proliferating macrophage (Fig. 7A and B). Focusing on the interaction between tumor cells and other cells, it is also found that there are various cell interactions in the immune active subtypes, including anti-tumor immune cells such as B cells, CD8⁺ T cells and NK cells. Like the immune resistance subtypes, the immune active subtypes also have the immunosuppressive interaction, which may be one of the reasons for the occurrence and development of tumors (Fig. 7C-D). Switching off immunosuppressive responses or boosting the action of anti-tumor immune cells could improve patient outcomes.

Then, we analyzed tumor-interacting pathways and found that the immunosuppressive pathways, such as MIF – (CD74 + CXCR4/CD44) and MDK-NCL, were enriched in two immune subtypes, which could serve as one of the potential therapeutic targets. In addition, a group of specific immune-related pathways (CXCL-CXCR) enriched in immune active subtypes between tumor cells and chemokine macrophages, which may be another potential target for clinical application (Fig. 7E-F).

In conclusion, the IRS scoring system still has a good ability to distinguish immune subtypes in single-cell level, which is helpful for clinical implementation of different treatment plans for patients with different immune subtypes, as well as the combination of immunotherapy and vaccine treatment.

4. Discussion

Immunotherapy and targeted therapy have brought enormous success to the treatment of LUAD. However, the combination therapy has not achieved satisfactory clinical results. One of the reasons is that targeted drugs rarely target the immune microenvironment. In this study, we identified two specific antigens ZC3H12D and TXNDC5 to develop an mRNA vaccine which could be recognized by APCs and presented to T cells and captured by B cells to activate an anti-tumor response, thus serving as an important complement to immunotherapy and provided a novel treatment method for LUAD patients. Although these candidates need further clinical validation, their potential for mRNA construction is supported by previous studies.

ZC3H12D, also referred to as MCP1P4, TFL and p34, has been considered a novel tumor suppressor gene in lung cancer [21,22]. ZC3H12D exerts crucial functions in immune modulation, and the deregulation of ZC3H12D may contribute to the onset of follicular lymphoma [23]. Tomita et al. pointed out that tumor-bearing lungs enhanced ZC3H12D expression in leukocytes in vitro co-culture system. ZC3H12D recruited to the cell membrane and captures the 3'-untranslated region of naked nonvesicular extracellular IL1 β -mRNA. Then, nex-IL1 β -mRNA was transported to the nucleus which increased anti-apoptotic gene expression, interferon- γ production and migration activity, resulting in the killing and anti-metastasis of cancer cells in mice [24]. TXNDC5, which is located in the endoplasmic reticulum and Golgi apparatus, plays a pivotal role in protein folding and anterograde transport. Importantly, TXNDC5 was found to have abnormally high expression in various tumors, such as lung cancer, prostate cancer, liver cancer, and esophageal squamous cell carcinoma [25,26]. Mo et al. reported that the markedly up-regulated TXNDC5 could promote cell growth, migration, and invasion of clear cell renal cell carcinoma and enhance the chemotherapy sensitivity of cells [27]. Similar findings were reported in the cervical tumor. The high expression of this gene contributes to abnormal angiogenesis, vasculogenic mimicry, and metastasis of cervical tumor cells by down-regulating SERPINF1 and TRAF1 expression [28]. Moreover, TXNDC5 can promote the secretion of IL-6 and IL-8 to promote the inflammatory response in rheumatoid arthritis [29]. IL-6 and IL-8, as members of the interleukin family, play an important role in immune responses such as T cell activation, so the effect of TXNDC5 on the tumor microenvironment needs further study [30,31]. These studies all support that ZC3H12D and TXNDC5 can be used as a biomarker for LUAD progression and activate a favorable immune response, and thus are candidates for LUAD vaccine development.

The subtyping criteria developed for LUAD could be well applied to finding the appropriate patient for vaccine treatment. This study subdivided LUAD patients into two immune subtypes based on immune checkpoint expression profiles. These two

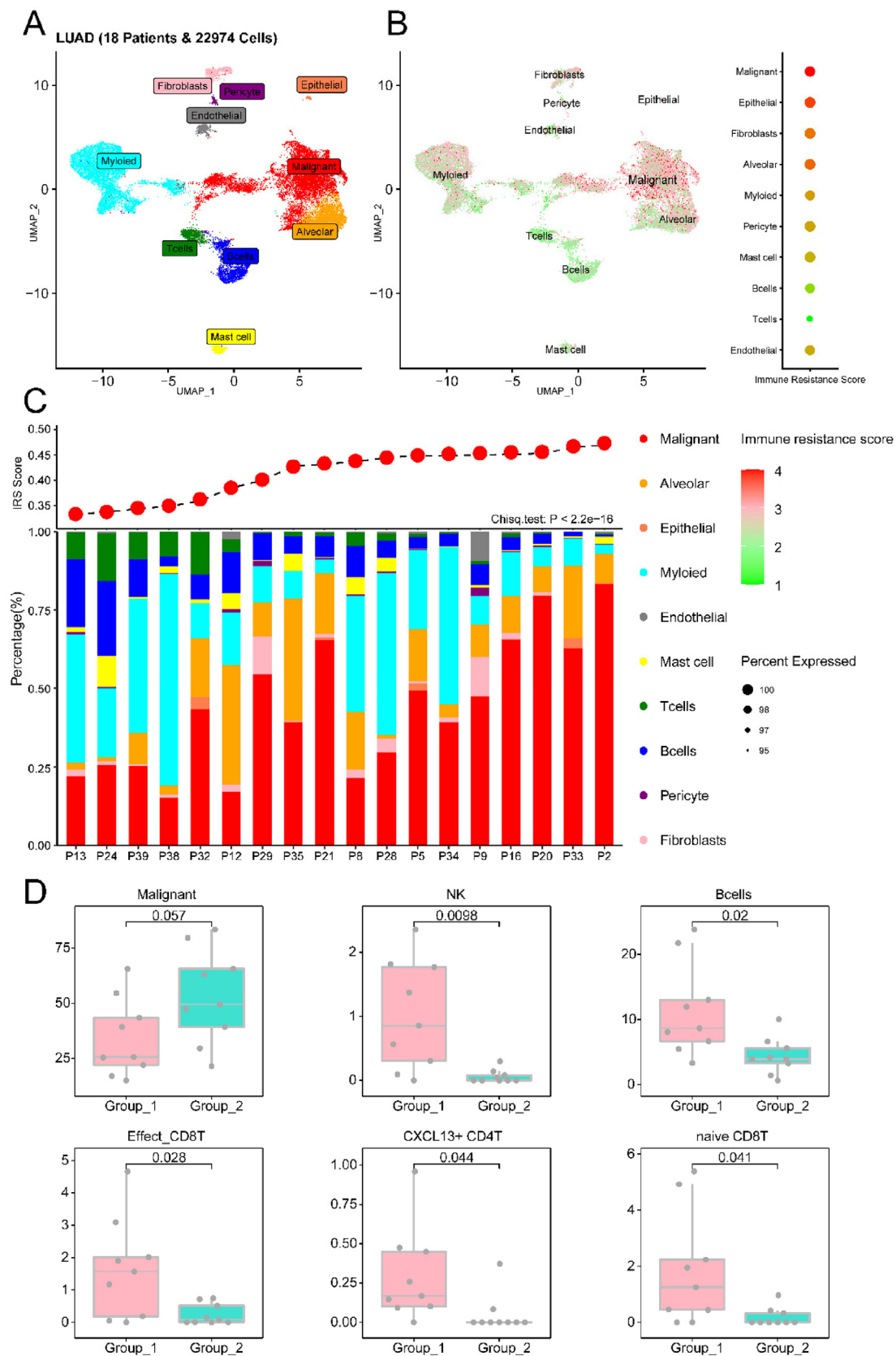


Fig. 6. Characteristics of single-cell profiling in different immune subtypes (A) UMAP plot show profiles of immune cells. (B) Expression of IRS among different types of immune cells. (C) The expression of IRS in different patients and the proportion of cells in the immune microenvironment. The scatter plot above shows the IRS levels of different patients. The histogram below shows the proportion of cells in the immune microenvironment of different patients (D) Differences in cell proportions between two immune subtypes.

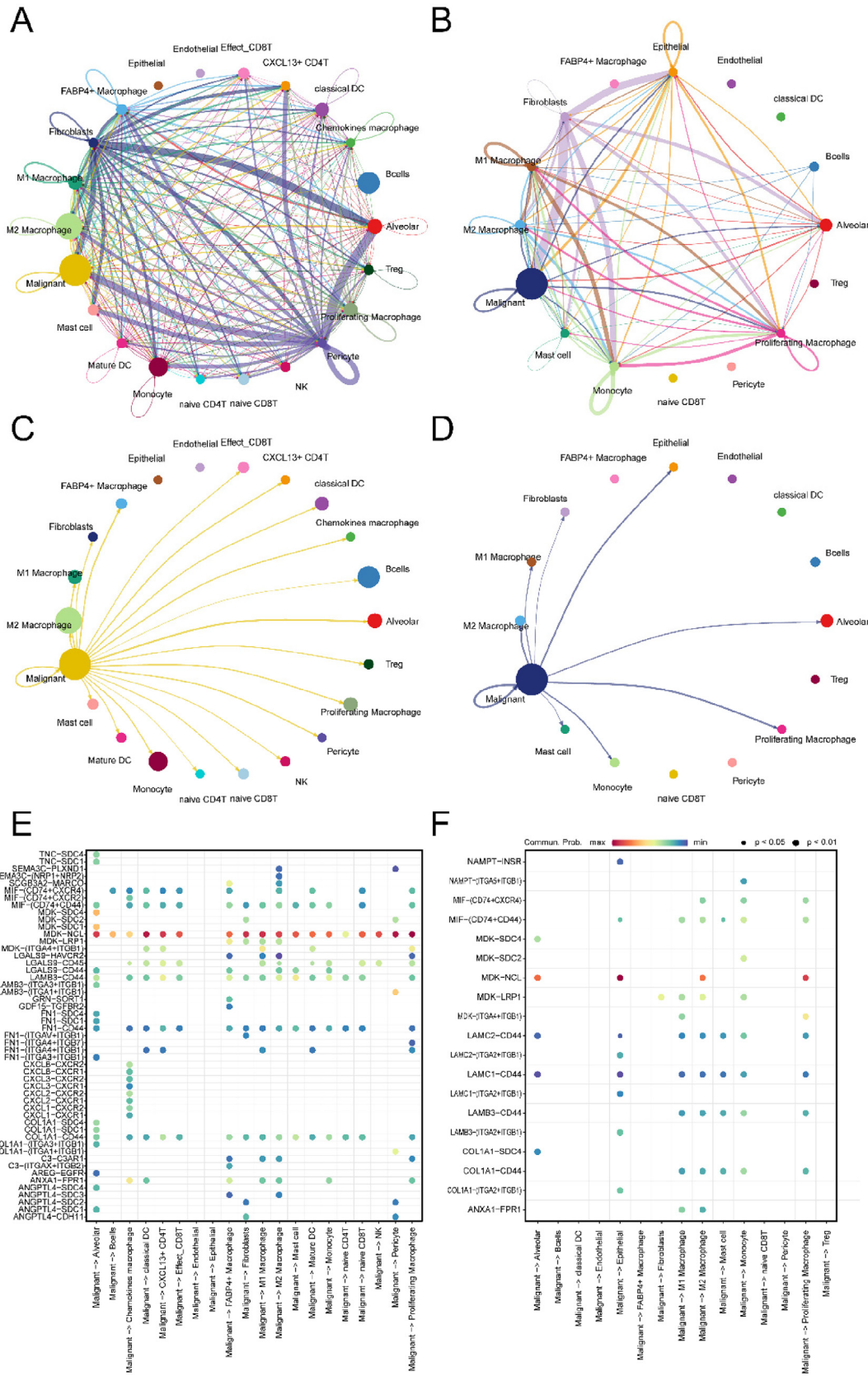


Fig. 7. Cell interactions in the immune microenvironment of different immune subtypes (A-B) The interaction of all cells in the immune microenvironment of the immune-active subtype (A) and the immune-resistant subtype (B). (C-D) Interaction between tumor cells and other cells in the immune microenvironment of the immune-active subtype (C) and the immune-resistant subtype (D). (E-F) Activated pathways of tumor cells interacting with other cells in the immune microenvironment of the immune-active subtype (E) and the immune-resistant subtype (F).

subtypes had prominent clinical, molecular, and cellular features, respectively. First, patients in group 1 showed a prolonged prognosis compared to group 2, which had a higher probability of tumor progression and recurrence. Subsequently, we found that multiple immune-related pathways were enriched in group 1, whereas tumor proliferative and metabolic pathways were enriched in group 2. Considering the tumor heterogeneity and complicated tumor immune microenvironment, which may lead to individual therapeutic responses with mRNA vaccines, we further explore the immune cell infiltration between the two groups. Compared to group 2, group 1 revealed significantly higher immune-related scores and most immune cells that were enriched in group 1, especially active B cells and CD4+/CD8+ T cells. According to the results, we identified group 1 as immune active subtype and group 2 as immune resistance subtype. Notably, ZC3H12D and TXNDC5 were observably increased in immune active subtype, which is consistent with our previous conjecture that the expression of both genes is associated with immune cell infiltration and immune response. Previous studies have supported our results that infiltration of CD8+ and CD4+ T cells was an independent prognostic factor in multiple malignancies [32]. Patients with high levels of CD8+ and CD4+ T cells had significantly higher survival rates than all other groups. In a study of colon cancer and immune infiltration, researchers found that an active immune response in tumors reduced early metastatic and invasion. Moreover, the increased immune cells infiltration predictor improved overall survival [33].

The above results show that our immune subtype classification is effective, and the immune resistant subtype can be classified as a “cold” tumor with less immune response and strong proliferation and metastasis ability, so it is not suitable for the application of immunotherapy and vaccines, while the immune active subtype can be classified as a “hot” tumor with a high immune response and immune cell infiltration, thus suitable for the application of the ZC3H12D and TXNDC5 vaccines we have discovered.

In the clinic, combination therapy is typical for LUAD patients, and different immune statuses need a customized regimen. Hence, we screened potential treatment targets between the two groups. Consistent with the result above, many clinically actionable genes were found in immune active subtype, while only a few proliferation-related genes were found in immune resistant subtype. Subsequently, the results of IRS-related drug analysis showed that 58 drugs were identified as IRS-sensitive drugs and thus suitable for immune resistant subtype. Based on the CTRP data, we identified vandetanib and erlotinib as low-IRS sensitive, while vincristine, doxorubicin, clofarabine, and etoposide as high-IRS sensitive. As a result, mRNA vaccines combined with suitable anti-tumor drugs may bring a new strategy to LUAD treatment.

Tumors are complex heterocellular systems containing epithelial cells, fibroblasts, and multiple immune cell types. The TME and cell–cell communication regulate cancer progression and influence therapeutic response [34]. Elucidating the cellular interactions in the microenvironment will help predict the response to drugs or vaccines. Therefore, we validated the IRS scoring system at the single-cell level to obtain a more pronounced intratumoral heterogeneity landscape. After calculating the IRS of each cell in GSE128071, it was found that the cells in the malignant cluster had significantly higher IRS scores, the proportion of tumor cells was significantly increased in the immune resistant subtypes, and anti-tumor immune cells were significantly concentrated in the immune active subtypes. Importantly, immune active subtypes had abundant cellular interactions between tumor cells and anti-tumor cells such as B cells, CD8+ T cells, and NK cells in TME. In contrast, immune resistant subtypes clearly lacked cellular interactions. This may be one of the reasons for the rapid progression of cold tumors. Furthermore, an

immune-related pathway, CXCL-CXCR, is enriched in immune-active subtypes, which could be the potential therapeutic targets for LUAD. Although we have demonstrated the reliability of our screening for tumor antigens and immunophenotypes from bulk sequencing and single-cell sequencing, further in vitro and in vivo experiments are required to validate these findings, including mRNA vaccine production, immune cell co-culture, and lung adenocarcinoma syngeneic model, etc.

In summary, we comprehensively assessed the immune microenvironment features of lung adenocarcinoma at the whole tissue and single-cell levels, thereby finding immune-related tumor-specific antigens for developing RNA vaccines and defining immune subtypes suitable for immunotherapy. Our study will help strengthen knowledge of TME characteristics and provides important guidance on immunotherapy regimens.

Data availability

The HTRNA-Seq, simple nucleotide variation, copy number variation data and related clinical information and single cell RNA sequencing data were described in method section “Data Extraction and Pre-processing”. All analyses were performed in R software. Specific content is explained in each section of Methods.

Funding

This work was supported by the National Natural Science Foundation of China (No.81974367, No.82003065, No.82172655), the National Natural Science Foundation of Hunan Province (No.2020JJ4132, No.2021JJ70023, 2019JJ50941), and National Multidisciplinary Cooperative Diagnosis and Treatment Capacity Building Project for Major Diseases (Lung Cancer, No. z027002).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank BioRender (biorender.com) for flowchart material.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2022.08.066>.

References

- [1] Siegel RL et al. *Cancer Statistics, 2021*. *CA Cancer J Clin* 2021;71(1):7–33.
- [2] Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. *Clin Chest Med* 2020;41(1):1–24.
- [3] Wang M, Herbst RS, Boshoff C. Toward personalized treatment approaches for non-small-cell lung cancer. *Nat Med* 2021;27(8):1345–56.
- [4] Reck M et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med* 2016;375(19):1823–33.
- [5] Carbone DP et al. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N Engl J Med* 2017;376(25):2415–26.
- [6] Duma N, Santana-Davila R, Molina JR. Non-small cell lung cancer: epidemiology, screening, diagnosis, and treatment. *Mayo Clin Proc* 2019;94(8):1623–40.
- [7] Sayour EJ, Mendez-Gomez HR, Mitchell DA. Cancer vaccine immunotherapy with RNA-loaded liposomes. *Int J Mol Sci* 2018;19(10).
- [8] Guan S, Rosenecker J. Nanotechnologies in delivery of mRNA therapeutics using nonviral vector-based delivery systems. *Gene Ther* 2017;24(3):133–43.
- [9] Pardi N et al. mRNA vaccines – a new era in vaccinology. *Nat Rev Drug Discov* 2018;17(4):261–79.

- [10] Sobhani N et al. Therapeutic cancer vaccines: From biological mechanisms and engineering to ongoing clinical trials. *Cancer Treat Rev* 2022;109:102429.
- [11] Hao Y et al. Integrated analysis of multimodal single-cell data. *Cell* 2021;184(13):3573–3587.e29.
- [12] Mayakonda A et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018;28(11):1747–56.
- [13] Hänzelmann S, Castelo R, Guinney J. GSV: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* 2013;14:7.
- [14] Zeng D et al. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front Immunol* 2021;12:687975.
- [15] Charoentong P et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* 2017;18(1):248–62.
- [16] Sjöstedt E et al. An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science* 2020;367(6482).
- [17] Auslander N et al. Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat Med* 2018;24(10):1545–9.
- [18] Afrache H et al. The butyrophilin (BTN) gene family: from milk fat to the regulation of the immune response. *Immunogenetics* 2012;64(11):781–94.
- [19] Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012;12(4):252–64.
- [20] Campbell KS, Purdy AK. Structure/function of human killer cell immunoglobulin-like receptors: lessons from polymorphisms, evolution, crystal structures and mutations. *Immunology* 2011;132(3):315–25.
- [21] Wakahashi S et al. Transformed Follicular Lymphoma (TFL) Predicts Outcome in Advanced Endometrial Cancer. *Cancer Epidemiol Biomarkers Prev* 2018;27(8):963–9.
- [22] Wang M et al. Identification of a novel tumor suppressor gene p34 on human chromosome 6q25.1. *Cancer Res* 2007;67(1):93–9.
- [23] Minagawa K et al. Dereglulation of a possible tumour suppressor gene, ZC3H12D, by translocation of IGH@ in transformed follicular lymphoma with t(2;6)(p12;q25). *Br J Haematol* 2007;139(1):161–3.
- [24] Tomita T et al. Extracellular mRNA transported to the nucleus exerts translation-independent function. *Nat Commun* 2021;12(1):3655.
- [25] Chawsheen HA et al. A critical role of the thioredoxin domain containing protein 5 (TXNDC5) in redox homeostasis and cancer development. *Genes Dis* 2018;5(4):312–22.
- [26] Wang X, Li H, Chang X. The role and mechanism of TXNDC5 in diseases. *Eur J Med Res* 2022;27(1):145.
- [27] Mo R et al. High TXNDC5 expression predicts poor prognosis in renal cell carcinoma. *Tumour Biol* 2016;37(7):9797–806.
- [28] Xu B et al. TXNDC5 is a cervical tumor susceptibility gene that stimulates cell migration, vasculogenic mimicry and angiogenesis by down-regulating SERPINF1 and TRAF1 expression. *Oncotarget* 2017;8(53):91009–24.
- [29] Lu Q et al. TXNDC5 protects synovial fibroblasts of rheumatoid arthritis from the detrimental effects of endoplasmic reticulum stress. *Intractable Rare Dis Res* 2020;9(1):23–9.
- [30] Tanaka T, Narazaki M, Kishimoto T. Interleukin (IL-6) Immunotherapy. *Cold Spring Harb Perspect Biol* 2018;10(8).
- [31] Waugh DJ, Wilson C. The interleukin-8 pathway in cancer. *Clin Cancer Res* 2008;14(21):6735–41.
- [32] Jochems C, Schlom J. Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity. *Exp Biol Med (Maywood)* 2011;236(5):567–79.
- [33] Pagès F et al. Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med* 2005;353(25):2654–66.
- [34] AlMusawi S, Ahmed M, Nateri AS. Understanding cell-cell communication and signaling in the colorectal cancer microenvironment. *Clin Transl Med* 2021;11(2):e308.