

Original Research

Development and Validation of a Prognosis-Prediction Signature for Patients with Lung Adenocarcinoma Based on 11 Telomere-Related Genes

Jia Liu¹, Sha Sha¹, Jian Wang¹, Xiaowei Gu¹, Menghua Du², Xu Lu^{2,*}¹Department of Radiotherapy, Jiangyin Hospital Affiliated to Nantong University, 214400 Jiangyin, Jiangsu, China²Department of Radiotherapy, Suzhou Ninth People's Hospital, 215200 Suzhou, Jiangsu, China*Correspondence: dirty.6@163.com (Xu Lu)

Academic Editor: Fu Wang

Submitted: 31 March 2023 Revised: 15 June 2023 Accepted: 21 June 2023 Published: 20 October 2023

Abstract

Background: The occurrence and progression of lung cancer are correlated with telomeres and telomerase. Telomere length is reduced in the majority of tumors, including lung cancers. Telomere length variations have been associated with lung cancer risk and may serve as therapeutic targets as well as predictive biomarkers for lung cancer. Nevertheless, the effects of telomere-associated genes on lung cancer prognosis have not been thoroughly studied. We aim to investigate the relationship between telomere-associated genes and lung cancer prognosis. **Methods:** The Cancer Genome Atlas and Genotype-Tissue Expression databases were used as training sets to build a predictive model. Three integrated Gene Expression Omnibus datasets served as validation sets. Using cluster consistency analysis and regression with the least absolute shrinkage and selection operator, we developed a telomere-related gene risk signature (TMGsig) based on 11 overall survival-related genes (*RBBP8*, *PLK1*, *DSG2*, *HOXA7*, *ANAPC4*, *CSNK1E*, *SYAP1*, *ALDOA*, *PHF1*, *MUTYH*, and *PGSI*). **Results:** The results indicated a negative outcome for the high-risk score group. Immunological microenvironment and somatic mutations differed between the high- and low-risk groups. A statistically significant difference existed between the low-risk and high-risk groups in terms of the expression levels of B cells and CD4 cells, and the risk score was essentially inversely linked with immune cell expression. **Conclusions:** TMGsig can predict outcomes in patients with lung adenocarcinoma.

Keywords: telomere; prognosis; signature; lung adenocarcinoma; TCGA; GEO

1. Introduction

Lung cancer is the most lethal malignant tumor, claiming over a million lives annually worldwide [1]. Different pathological subtypes of lung cancer carry different prognoses and clinicopathological features. Lung cancer has a 5-year survival rate of only approximately 18%, and survival rates for different stages of lung cancer vary greatly [2]. Due to the high heterogeneity of lung cancer, and despite the availability of numerous treatment methods, the effectiveness of most clinical medications for treating the condition remains relatively low. Chemotherapy, radiotherapy, targeted therapy, and even immunotherapy may be effective, but many tumors eventually develop drug resistance, which complicates the clinical treatment of lung cancer. Its poor survival rate and complex treatment indicate a key conclusion: the recovery and survival rates of patients can only be successfully increased by early identification, early diagnosis, and early treatment. In this sense, multi-gene models are generally better at predicting tumors [3]. As a result, it is possible to more precisely pinpoint patient subgroups with poor survivals, to deliver more precise adjuvant therapy and follow-up. Decision-making and individualized care are also aided by this. During cell division, the DNA near the ends of chromosomes is gradually eroded

due to the incapacity of chromosomal DNA replication processes to effectively copy this DNA. A reverse transcriptase called telomerase adds telomere sequences to the ends of chromosomes, making up for the DNA loss that happens in the region during cell division. The telomere lengths of somatic cells gradually decrease during their life cycles because most normal tissues lack telomerase activity. However, telomerase is reactivated in 85–90% of human tumors [4]. In 2009, for discovering the processes by which telomeres and telomerase protect chromosomes, three scientists received the Nobel Prize in Physiology or Medicine. Since then, research on telomeres and telomerase in cancer has gradually gained interest, particularly the blocking of the telomere elongation mechanism, and telomerase inhibitors as a potential new therapeutic target for anti-tumor therapy [5].

Telomere shortening may be a risk factor or predictor of susceptibility to lung cancer, according to some research [6]. However, one prospective observational study of heavy smokers indicated that lung cancer risk was higher in people with longer telomeres [7]. Similar conclusions were reached in a follow-up study of 26,540 Chinese Singaporeans [8]: a Cox regression model was used, and the results revealed that individuals in the top 20% concerning telomere length had an increased risk of lung adenocarcinoma



Table 1. Including cohorts in our study.

Cohort names	RNA Type	Cancer type	Number of patients	Cohort Type
TCGA-LUAD	RNA-seq	LUAD	575	Training
GSE70249	Microarray	LUAD	398	Validation
GSE31210	Microarray	LUAD	226	Validation
GSE11969	Microarray	LUAD	149	Validation

TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma.

(LUAD) compared with those in the bottom 20% (hazard ratio [HR] = 2.84; 95% confidence interval [CI]: 1.94–4.14; $p < 0.0001$). The likelihood of acquiring LUAD or other disease types, however, did not correlate with telomere length [8]. It has been suggested that studies on telomere length and lung cancer risk, especially concerning the underlying mechanism, require further development. Regarding telomere length and the prognosis of lung cancer patients, studies have shown that telomere lengths in peripheral white blood cells and chances of mortality are correlated in patients with cancer [9–11]. Although the results of different studies on the matter have not been entirely consistent, the link between telomeres and lung cancer merits further investigation.

Given the link between telomeres and lung cancer, we created a predictive model for LUAD based on telomere-associated genes. We also discussed the relationship between this model and the immune micro-landscape. Hopefully, this strategy may support physicians in making crucial treatment-related decisions in the future.

2. Materials and Methods

2.1 Acquisition of Data

The Sangerbox online platform (<http://vip.sangerbox.com/>) was used to access the Genotype-Tissue Expression (GTEx) database and The Cancer Genome Atlas (TCGA) database to view clinical and gene expression data for patients with LUAD [12]. The datasets for GSE31210, GSE11969, and GSE72049 were retrieved from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>). Table 1 provides an overview of the clinical characteristics of the four groups. Furthermore, the TelNet database (<http://www.cancertelsys.org/telnet/>) was used to acquire telomere-associated genes [13].

2.2 Functional Enrichment Analysis and Differentially Expressed Gene Screening

The TCGA dataset provided RNAseq results and associated clinical data for 516 LUAD tumors. The Limma package for R (version 3.40.2, R Foundation for Statistical Computing, Vienna, Austria) was used to investigate differential mRNA expression. Adjusted p -values were examined in TCGA or GTEx to correct false positive results, and “Adjusted $p < 0.05$ with $|\log_2FC| > 1$ ” was used as a cutoff for Differentially Expressed Genes (DEGs).

The data was analyzed through functional enrichment analysis. To better understand the oncogenic activities of target genes, the ClusterProfiler package (version: 3.18.0) for R was used to analyze the Gene Ontology (GO) functions of potential targets and enrichments identified via the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway.

2.3 Cluster Consistency Analysis Based on DEGs

The intersection between DEGs and telomere-related genes yielded Telomere related DEGs (TM-DEGs). To investigate the relationships between the expression levels of TM-DEGs and LUAD subtypes, ConsensusClusterPlus was used to perform cluster analysis [14] on all 514 LUAD patients in the TCGA dataset. After 10 iterations of agglomerative pam clustering with a 1-Pearson correlation distance, 80% of the data were resampled. The ideal number of clusters was established using the empirical cumulative distribution function plot. As a result, prognostic analysis was performed on the two identified groups.

2.4 The Creation and Verification of the Telomere-Related Genes Signature (TMGsig)

A telomere-related gene signature (TMGsig) was established using the least absolute shrinkage and selection operator (LASSO)-Cox regression analysis. The maxstat package for R was used to determine the optimal threshold of risk score, to divide each cohort into high-risk and low-risk groups. The survival analysis tool of the R software package was used to determine the prognostic difference between the two groups. The log-rank test was used to assess the substantial prognostic variances between the groups. A receiver operating characteristic (ROC) curve was used to evaluate the predictive ability of the signature. Finally, using risk scores based on the TCGA cohort, a nomogram was developed to forecast 1-, 3-, and 5-year survival.

2.5 Functional Enrichment Analysis of 11 Telomere-Related Prognostic Genes

We used the KEGG and GO enrichment methods to analyze 11 identified telomere-related prognostic genes.

Minimum and maximum gene sets were 5 and 5000, respectively, and statistical significance was defined as $p < 0.05$ and False Discovery Rate (FDR) < 0.1 .

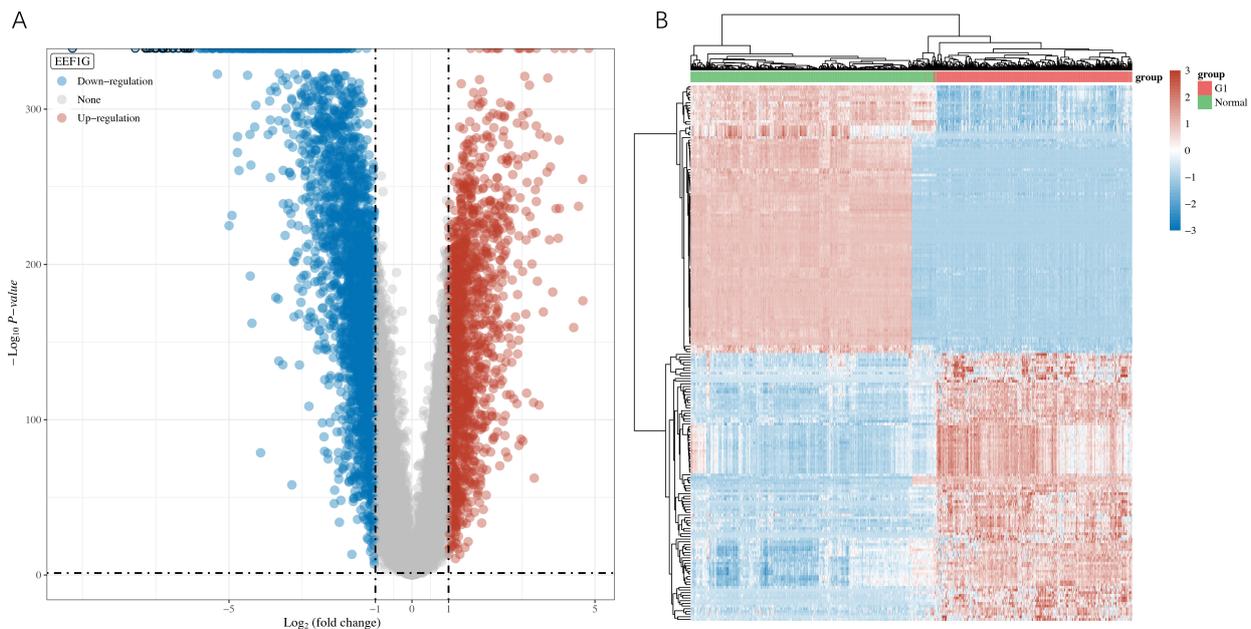


Fig. 1. Screening of Differentially Expressed Genes (DEGs). (A) Using fold change and adjusted p -values to construct a volcano map. (B) Heat map of differential gene expression.

2.7 Immune Correlation Analysis

The TCGA-LUAD cohort was used for Spearman correlation analysis between model score and immune score using the Tumor Immune Estimation Resource (TIMER) and quanTIseq algorithms, respectively, and with $p < 0.05$ considered statistically significant. All analyses were conducted in R v4.0.3.

2.7 Two Categories of Somatic Mutations and Tumor Microenvironment (TME) Cell Infiltration

The TCGA genomic data commons (GDC) data portal contained somatic mutation information in the Mutation Annotation Format (MAF) format, for LUAD samples. For waterfall mapping, the “Maftools” package in R software was used, which made it easier to visualize and summarize mutant genes and create a waterfall map of somatic mutations for high- and low-risk subgroups.

Using the TIMER and quanTIseq methods, we compared the immune microenvironment infiltration of the two risk subgroups. The findings were displayed as a landscape map.

3. Results

3.1 Identification and Functional Enrichment Analysis of DEGs

The differential gene analysis of cancer and standard tissue samples from the TCGA-LUAD cohort, compared with normal tissue samples from GTEx database revealed a total of 4269 DEGs, 1475 of which showed upregulation and 2794 of which showed downregulation (Fig. 1).

GO analysis revealed that the majority of up-regulated genes were primarily involved in organelle fission, nuclear division, and nuclear chromosome segregation. The majority of the down-regulated genes were involved in second messenger-mediated signaling, regulation of vasculature development, and neutrophil degranulation neutrophil activation during the immune response.

According to our KEGG analysis, up-regulated genes are mainly enriched within the functional regions of the p53 signaling pathway, tight junction, and cell cycle. However, the active regions of the PI3K/Akt signaling pathway, the Ras/Rap1 signaling pathway, and the MAPK signaling pathway were enriched with down-regulated genes (Fig. 2).

3.2 TM-DEGs Acquisition and Co-Cluster Identification

A total of 2089 telomere-related genes were obtained from the previous website. The intersection of DEGs and telomere-related genes was then used to identify TM-DEGs (488 genes, among which 186 were up-regulated and 302 were down-regulated). Three NA expression values were eliminated, leaving 485 remaining genes. After dividing the patients into C1 and C2 groups, a separate survival analysis was conducted on each group. With a p value of 2.4×10^{-3} , the Kaplan-Meie (KM) curve demonstrated a statistically significant difference in overall survival (OS) between the two groups (Fig. 3A,B).

3.3 A Prognostic Signature Based on 11 TM-DEGs

Data on expression levels of TM-DEGs, survival time, and survival state were combined using the R software tool glmnet. The Lasso-Cox approach was used to create a prognostic signature (Fig. 4A–E).

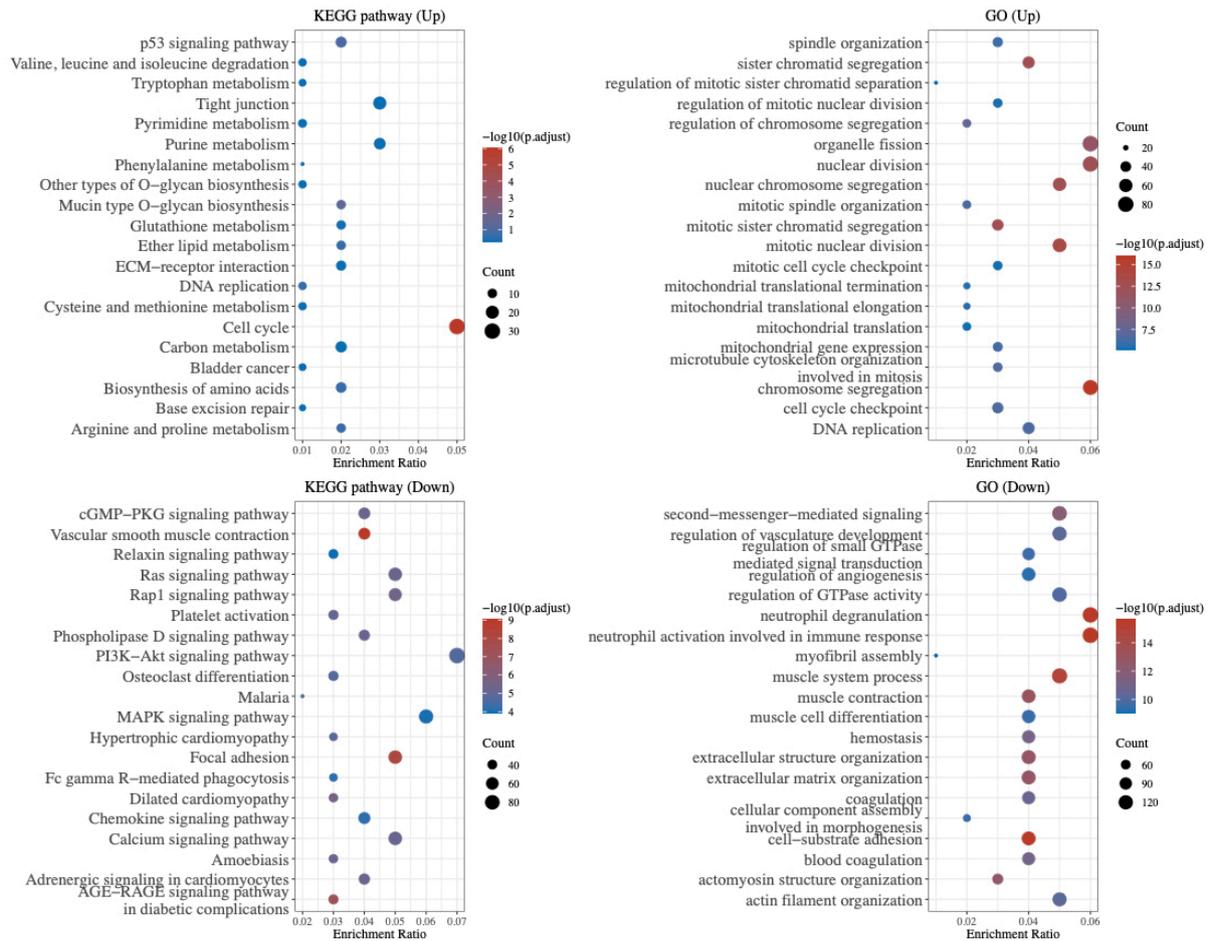


Fig. 2. Functional enrichment analysis of DEGs.

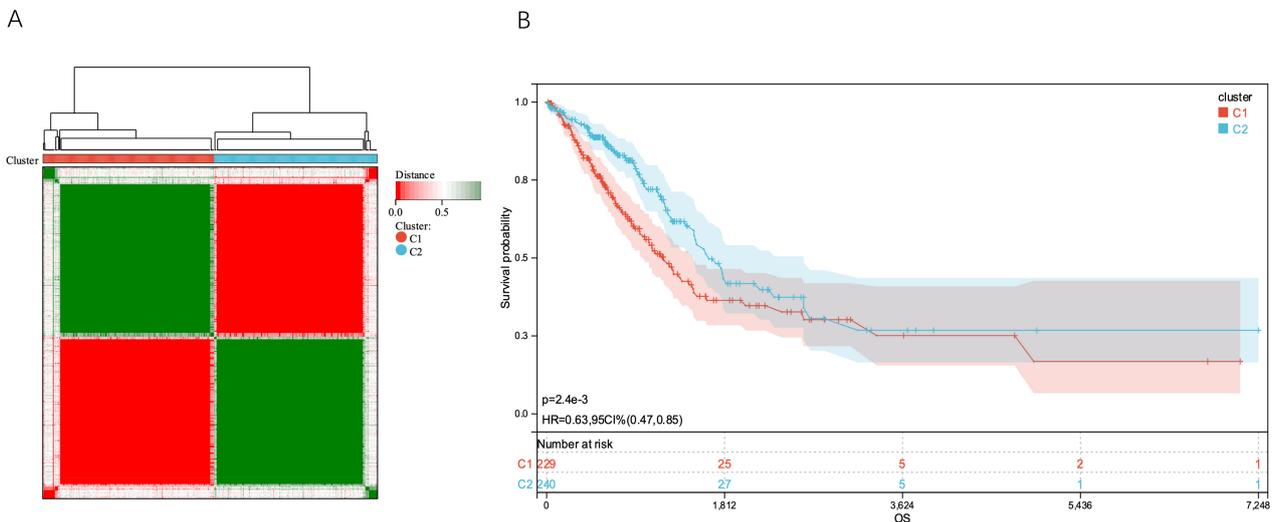


Fig. 3. Tumor classification based on TM-DEGs. (A) 514 LUAD patients were divided into two groups according to the uniform cluster matrix ($k = 2$). (B) Kaplan - Meier OS curves for the two clusters. LUAD, lung adenocarcinoma; OS, overall survival.

We set the Lambda value to 0.0617426400041608, and obtained the following model formula for the construction based on the 11 identified genes:

$$\begin{aligned} \text{riskscore} = & 0.1547 \times \text{HOXA7}_{\text{ss}} - 0.0301 \times \text{MUTYH} \\ & - 0.1130 \times \text{ANAPC4} - 0.0644 \times \text{PGS1} - 0.0260 \times \text{PHF1} + \\ & 0.1397 \times \text{PLK1} + 0.0364 \times \text{SYAP1} + 0.0532 \times \text{CSNK1E} \\ & - 0.0004 \times \text{RBBP8} + 0.0692 \times \text{DSG2} + 0.0236 \times \text{ALDOA} \end{aligned}$$

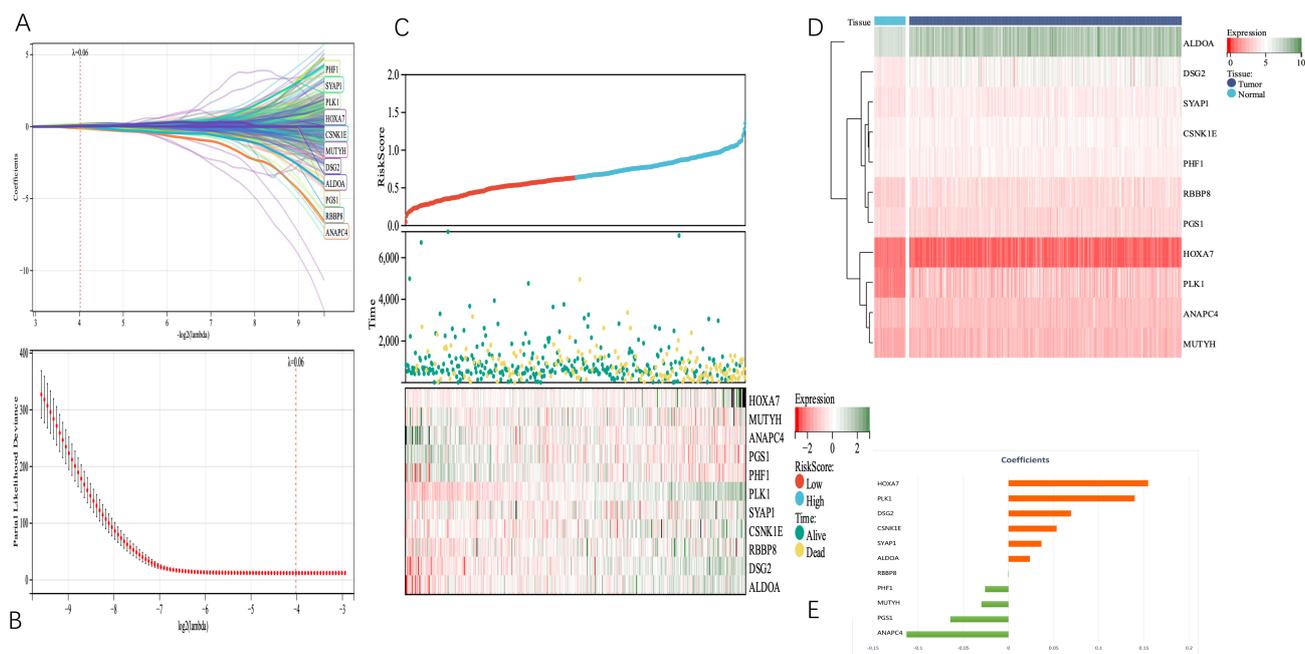


Fig. 4. Telomerase-related genes associated with tumor survival in LUAD patients were identified from the TCGA cohort. (A) and (B) 11 TM genes associated with OS were identified by LASSO regression and cross-validation. (C) Distribution of telomerase-associated gene risk score (TMGsig), survival status, and 11 prognostic TMG expression levels in LUAD patients in the TCGA cohort. (D) The heat map showed the expression of 11 TM genes in tumor and normal tissues. (E) Coefficients of the 11 screened telomerase-related genes (TMGs) in the Lasso Cox regression model. LASSO, least absolute shrinkage and selection operator; TCGA, The Cancer Genome Atlas.

Additionally, the Cox method was used to examine the prognostic significance of TM-DEGs in 469 samples. These findings suggest that the risk score and Tumor-Node-Metastasis (TNM) stage were independent prognostic factors (Fig. 5).

Based on this, the patients were divided into high- and low-risk groups, and we discovered a statistically significant difference in prognosis ($p = 1.1 \times 10^{-13}$), indicating that individuals in the high-risk group had worse prognoses (Fig. 6A). After dividing the samples into two groups based on the risk score, we analyzed the cluster recognition results of patients in the high-risk group and discovered that the C1 cluster was the most prevalent. By contrast, the C2 cluster was more prevalent in the low-risk group (Fig. 6B).

ROC analyses were conducted at time points 365, 1095, and 1825, yielding area under the curve (AUC) values of 0.71, 0.70, and 0.70, respectively (Fig. 7).

By combining survival time data, survival status, age, TNM stage, and risk score, we created a nomogram with Cox methods and the R software package rms to assess the predictive value of these variables in 446 samples. The overall C-index of the model was 0.7097, 95% CI (0.6647–0.7548), $p = 7.0453 \times 10^{-20}$ (Fig. 8A,B).

Eleven TMGs were found to be mainly enriched in the nucleoplasm, cellular response to DNA damage, catalytic activity, and acting on DNA functional regions, as determined by GO analysis. KEGG analysis revealed that they

were primarily enriched in progesterone-mediated oocyte maturation, cell cycle, and oocyte meiosis (Fig. 9A,B).

3.4 Validation of the Prognostic Models Using Three External Cohorts from the GEO Database

For external validation, we used three cohorts from the GEO database to further validate the predictive value of the prediction signature. In the GSE31210 cohort of 226 patients with stage I-II LUAD, individuals at high risk had significantly shorter OS ($p = 2.0 \times 10^{-3}$) based on the KM curve. At 1, 3, and 5 years, the AUC values of the prediction model were 0.73, 0.59, and 0.69, respectively. The KM curves of GSE11969 ($n = 149$) and GSE72094 ($n = 398$) yielded identical results, ($p = 1.8 \times 10^{-3}$) and ($p = 1.0 \times 10^{-7}$), respectively. Meanwhile, the AUC values at 1, 3, and 5 years were 0.69, 0.57, and 0.56, respectively, for the GSE11969 cohort and 0.65, 0.63, and 0.62, respectively, for the GSE72094 cohort (Fig. 10A–C).

3.5 Somatic Mutations in Two Subgroups

We discovered that the cancer subtypes had distinct somatic mutant spectra. Although the most prevalent mutation types were consistent between the two groups, the relative frequencies differed between the different subtypes. The most common mutations in the high-risk group were *TP53*, *TTN*, *CSMD3*, *RYR2*, and *MUC16*, accounting for 62.5%, 62.6%, 50.4%, 48.7%, and 47.8% of the total num-

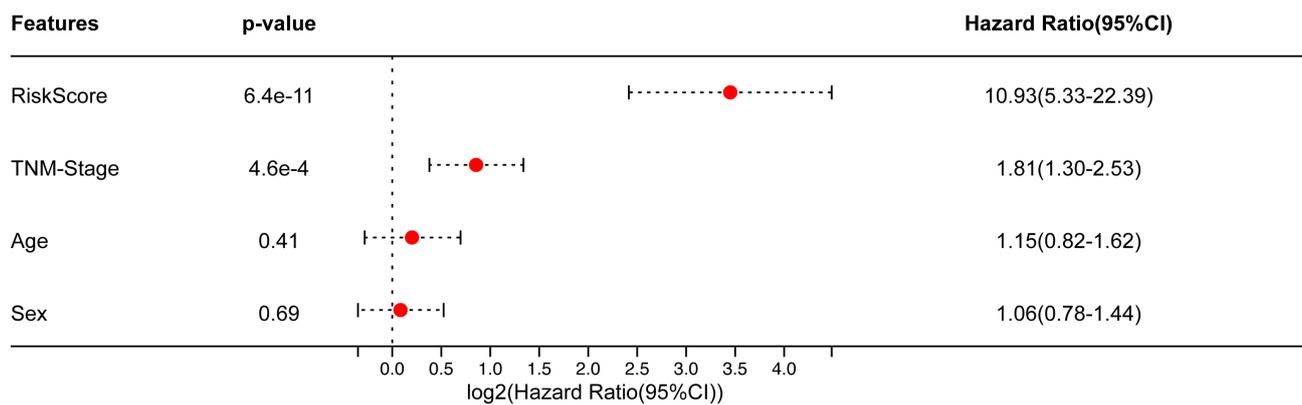


Fig. 5. In the TCGA-LUAD cohort, multivariate Cox results.

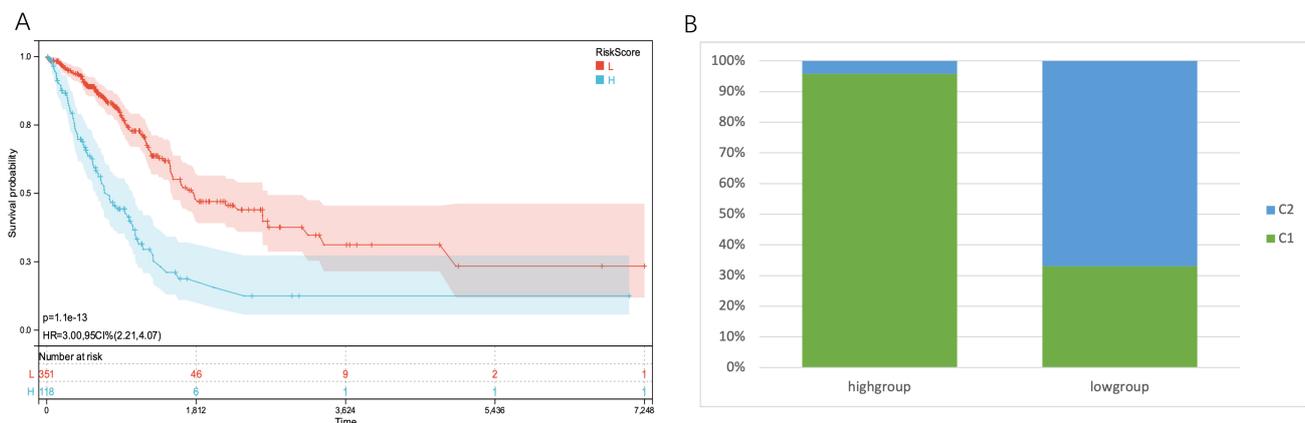


Fig. 6. KM curves of the two subgroups and the proportion of subtypes identified by clustering in the two groups. (A) In the TCGA-LUAD cohort, the KM curves of patients with high TMGsig risk scores and patients with low TMGsig showed apparent differences, and patients with high-risk score had a poorer prognosis. p -value = 1.1×10^{-13} . (B) The proportion of patients with C1 and C2 clusters in the different risk score subgroups.

ber of mutations, respectively. In the low-risk group, the most common mutations were *TP53*, *TTN*, *MUC16*, *CSMD3*, and *RYR2*, accounting for 47.3%, 45.3%, 39.3%, 35.9%, and 35.2% of the total number of mutations, respectively (Fig. 11A,B).

3.6 Relationship Between TMGsig and TME

We then investigated the connection between TMGsig and the tumor microenvironment (TME). The TIMER algorithm revealed a negative correlation between activated CD8 cells, active CD4 cells, B cells, and myeloid dendritic cells in patients with high-risk scores. The quanTiseq algorithm revealed that high-risk scores were associated with B cells, NK cells, macrophages, activated CD8 cells, and Treg, but not with active CD4 cells (Fig. 12A,B).

TME cell invasion is crucial to both carcinogenesis and treatment response. We looked at the frequency of several immune cell types in the high- and low-risk subgroups of the TCGA dataset to evaluate whether telomere-associated genes played major roles in the development of TMEs. Using TIMER and quanTiseq analysis of immune

cell types, we found that there were significant distinctions between the two groups in terms of the characteristics of TME cell invasion. For example, we discovered statistically significant differences between B cells and CD4 cells, indicating higher expression levels in the low-risk group (Fig. 13A–F).

4. Discussion

According to cancer statistics from 2021, lung cancer is the leading cause of cancer-related mortality in both men and women, accounting for over a quarter of all cancer-related deaths [15]. Low lung cancer survival rates are a result of a high percentage of patients (57%) having metastatic diseases by the time they are diagnosed.

When it comes to the development and prognosis of lung cancer, telomeres play a significant role. Using a publicly available database, we developed and validated a predictive signature, and then matched it to tumor immune microenvironments and somatic mutations. The genes in the signature serve a variety of functions in disease states. The Homeobox (*HOX*) gene family is closely associated with

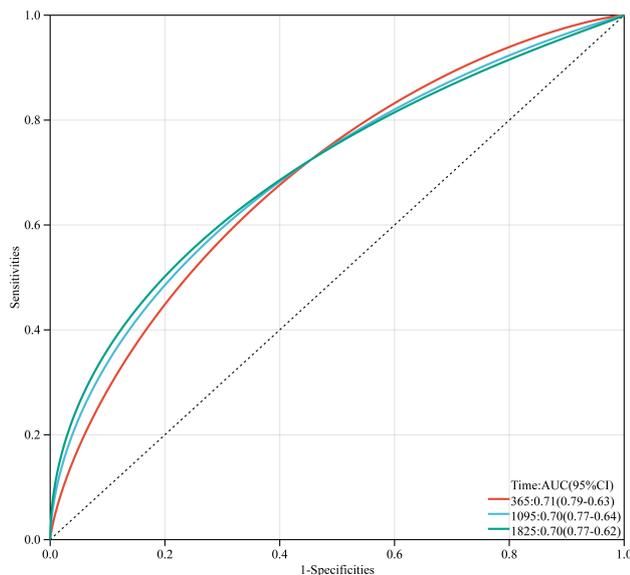


Fig. 7. Time-dependent receiver operator characteristic curves of TMGsig in the TCGA training cohort.

various cellular processes related to homeostasis, as well as tissue repair [16,17]. Several studies have shown that aberrant expression of the *HOX* gene is associated with the development of non-small cell lung cancer (NSCLC), as well as abnormal histological variants of adenocarcinoma and squamous cell carcinoma [18]. In one study on liver cancer, the *HOXA7* gene, which is part of the A cluster of chromosome 7, significantly enhanced proliferation, migration, and invasion *in vitro*, as well as tumor growth and metastasis [19]. This study found that the gene contributed the most as a risk factor to risk score, indicating that a worse prognosis was linked to its increased expression. Polo-like kinase 1 (*PLK1*) is the best-studied of the four Polo-like kinase subtypes that are essential for mitosis. *PLK1* expression is increased in various tumor cell lines, and is frequently associated with a poor prognosis [20]. An increase in *PLK1* activity leads to polyploid production, which is related to the growth of tumors. As a member of the classical cadherin family, Desmoglein 2 (*DSG2*) helps regulate intercellular connectivity and promotes the assembly of desmosomes [21]. Desmosomes are connected to cell adhesion junctions, which are crucial for the growth and invasion of cancer cells during tumor formation [22]. Studies have shown that the knockout of the *DSG2* gene can inhibit the proliferation of NSCLC [23]. Casein kinase 1 epsilon (*CSNK1E*) is a circadian gene that encodes DNA replication and repair-related members of the serine/threonine protein kinases and casein kinase I protein families. However, in normal cells, its primary function is to phosphorylate members of the Period family of circadian proteins, thereby regulating the average rhythm of cell growth and development. IC261, which has been reported to be a *CSNK1E* inhibitor in both *in vitro* and *in vivo* studies, also inhibits microtubule polymerization [24], which may lead to increased

cancer cell death [25]. One study suggested that the synergistic lethality between Tumor protein p53 (*TP53*) and *CSNK1E* is essential for the destruction of cancer cells [26]. Both activated WNT signaling [27] and *TP53* inactivation [28] are often associated with the development of cancerous tumors. Because *TP53* is mutated in 50% of all cancers, *CSNK1E* may be a promising drug target for patients with *TP53* mutations. Currently, the majority of Synapse associated protein 1 (*SYAP1*) research focuses on nervous system diseases. There is evidence that the DNA methylation and expression profiles of this gene differ between male and female glioma patients [29]. As a member of the aldolase family, *ALDOA* is highly expressed in various tissues such as gastric cancer [30], NSCLC [31], breast cancer [32], and oral squamous cell carcinoma [33]. As an independent prognostic factor, it is indicative of a poor prognosis [34,35]. In lung cancer, Aldolase, fructose-bisphosphate A (*ALDOA*) promotes tumor proliferation and is closely related to HIF-1 α , which can activate the HIF-1 α /MMP9 signal axis by blocking prolyl hydroxylase activity and promoting the invasion and metastasis of NSCLC [36]. It can encourage proliferation and G/S transition in NSCLC by activating the EGFR/MAPK signaling pathway [37]. Chang *et al.* [38] constructed an *ALDOA* mutant with an inhibited enzymatic activity. They found that *ALDOA* could still promote the transcription of *Oct4* by inhibiting the expression of miR-145, thus inducing dry lung cancer cells and leading to chemoresistant tumors. These results indicate that *ALDOA* affects tumor metabolism through both enzymatic and non-enzymatic processes. Studies have found that *sirt6* promotes DNA-end excision after DNA damage by deacetylation of *RBBP8*. This demonstrates that *SIRT6*-mediated *RBBP8* deacetylation can significantly increase cancer incidence and enhance the efficacy of the targeted therapy Gefitinib. *RBBP8* is thought to function similarly to *BRCA1.20* and *21* as a tumor suppressor [39]. As an essential member of the PcG protein family, PHD finger protein 1 (*PHF1*) interacts with various proteins and core complexes, including Ku70-Ku80 and radiation sensitive 50 genes (radiation sensitive 50, Rad50), structural maintenance of chromosomes protein 1 (*SMC1*), RNA helicase A (*RHA*, *RHA*, and *SMC1*—also known as *DHX9*), p53, and polycomb repressive complex 2 (*PRC2*), which are involved in DNA damage repair [40] and/or mediated the dormancy or apoptosis of damaged and cancerous cells [41]. *PHF1* is involved in complex biological processes that may mediate the balance of reconstruction following cell injury, and its effect on tumors cannot be defined simply by high or low expression levels. According to some research, *PHF1* is expressed weakly in pure non-small cell adenocarcinoma tissues but significantly in pure non-small cell squamous cell carcinoma tissues. It contributes to the proliferation, invasion, and death of non-small cell lung squamous cell carcinoma and adenocarcinoma cells, respectively, as an oncogene and an oncosuppressor gene, revealing its dual

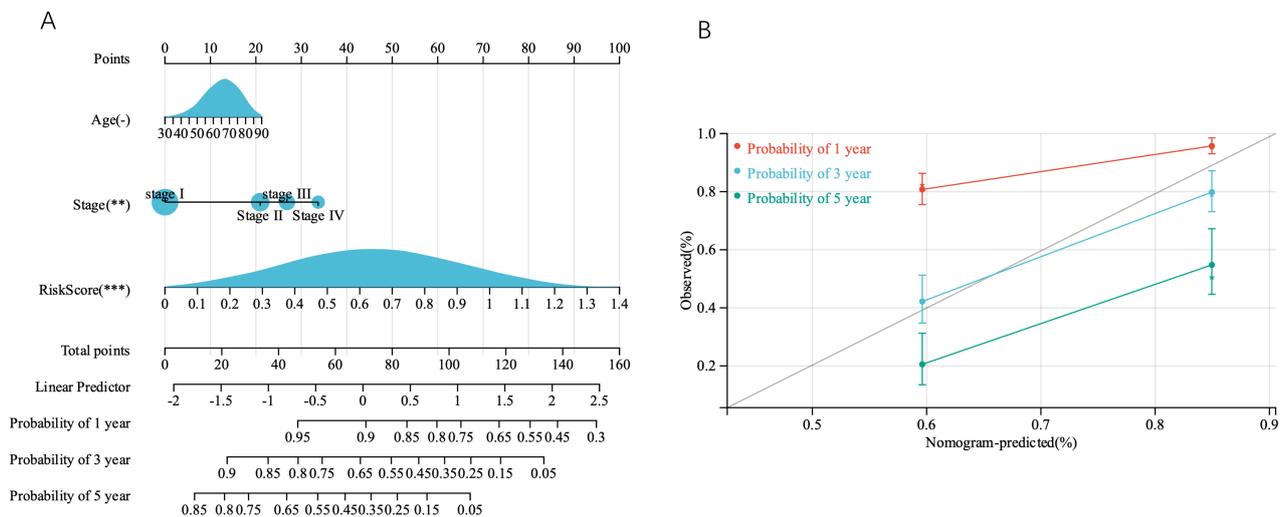


Fig. 8. Construction and validation of the nomogram. (A) A prognostic nomogram based on the risk score derived from the prognostic TMGsig and the age TNMstage, to predict 1-, 3-, and 5-year survival in patients with LUAD. (B) Calibration curves adjusted for discrepancies between the predicted and observed 1-, 3-, and 5- years survival rates.

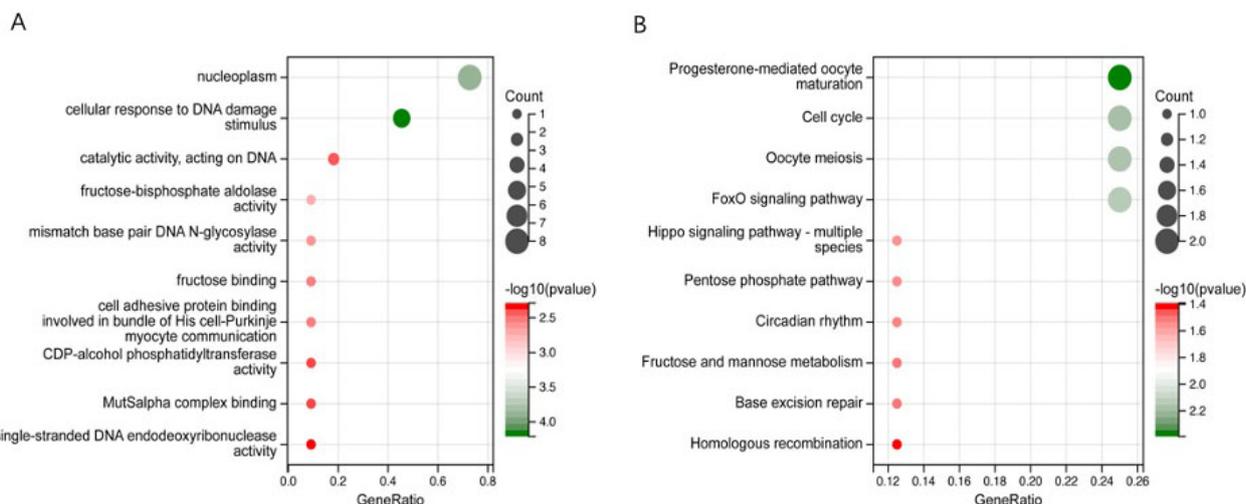


Fig. 9. Functional enrichment analysis of 11 TMGs. (A) Gene Ontology (GO) signaling pathway enrichment analysis. (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathway enrichment analysis.

structural properties [42]. A DNA glycosylase called MutY DNA glycosylase (*MUTYH*) is involved in the correction of mistakes after double-stranded DNA replication [43]. The molecular mechanism of *MUTYH* mutation in the development of colorectal cancer and *MUTYH*-associated polyposis has been studied extensively [44–46]. Current lung cancer research is centered on *MUTYH* single-nucleotide polymorphisms. There is evidence that the common *MUTYH* variant p.Gln324His may raise the risk of lung cancer because this affects its enzymatic activity similarly to the known cancer variant p.Gly396Asp, according to an *in vitro* study of the gene [47]. Phosphatidyl-glycerophosphate synthase 1 (*PGS1*) and Anaphase promoting complex subunit 4 (*ANAPC4*) have not been extensively studied in cancer;

however, given its analogous position in the biology of yeast telomeres, *ANAPC4* is predicted to influence telomere maintenance [48]. Further research must be carried out to determine the precise tools and purposes of these risk model genes.

This study divided TM-DEGs in LUAD patients into two clusters, by identifying cluster consistencies. Our predictive analysis showed that the two patient clusters had statistically different OSs, with cluster C1 patients having a shorter mean OS time ($p = 2.4 \times 10^{-3}$). Using LASOS-Cox, a survival time-related prognostic signature was then constructed. The OS time between the two prognostic groups was significantly different, indicating that patients in the high-risk group had shorter OSs ($p = 1.1 \times$

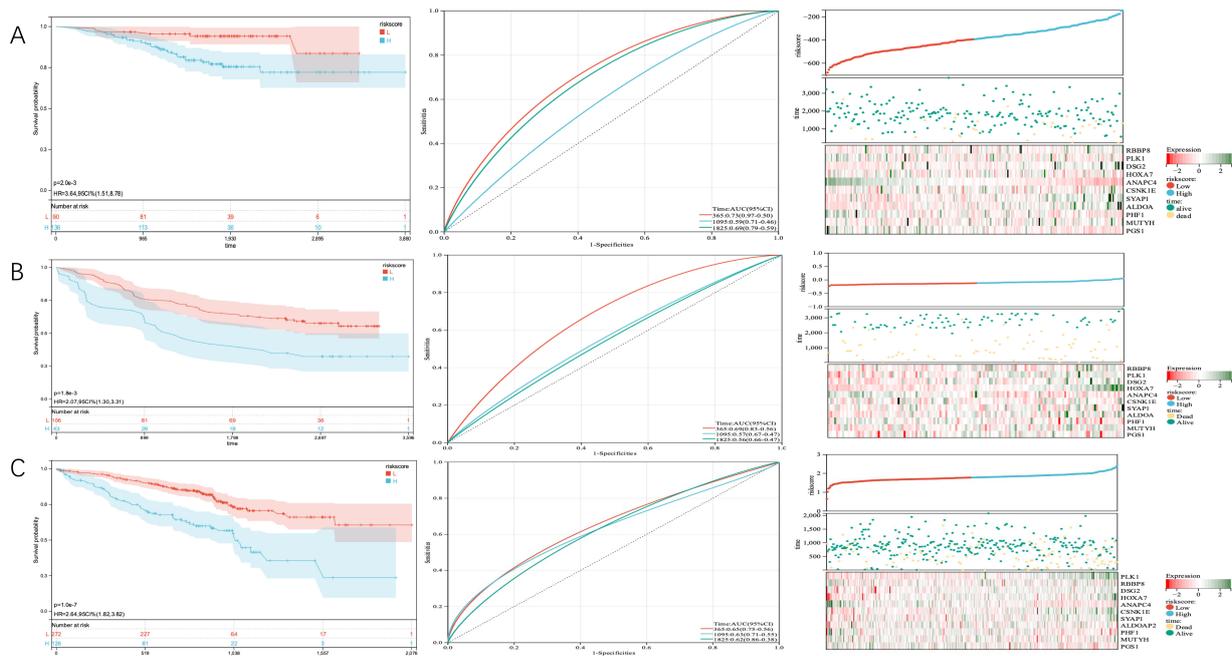


Fig. 10. KM curves of OS according to TMGSig, distribution plots of patients, and time-dependent ROC curves of TMGSig in the GEO validation cohort. (A) The GSE31210 cohort. (B) The GSE11969 cohort. (C) The GSE70249 cohort. ROC, receiver operating characteristic; GEO, Gene Expression Omnibus.

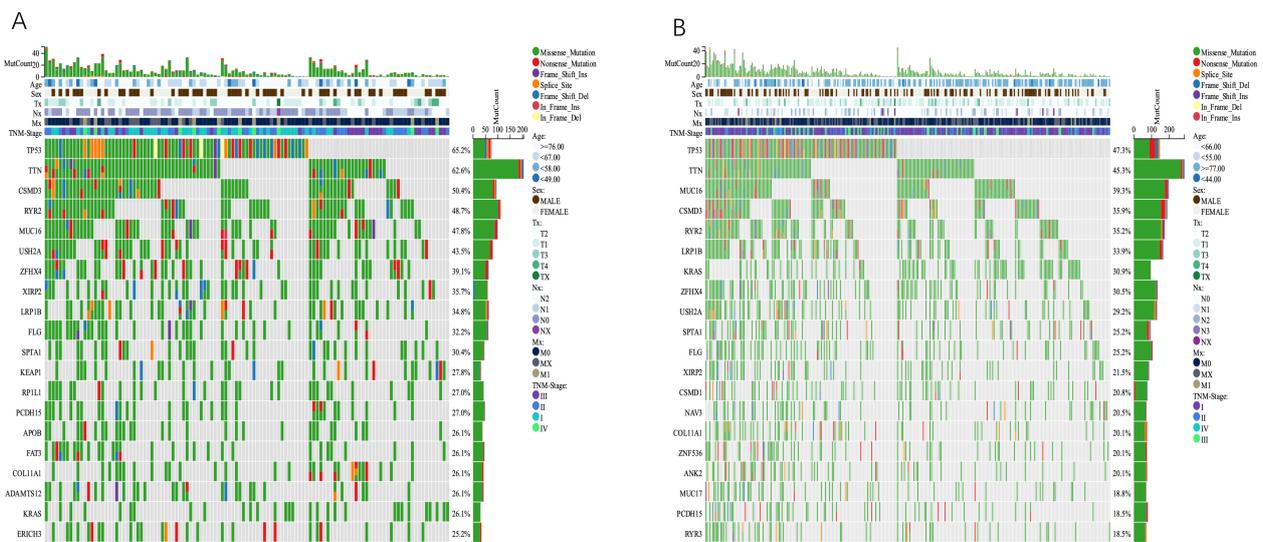


Fig. 11. Comparison of somatic mutations between different TMGSig subtypes. (A) Ocoprint representation of the top 20 most frequently mutated genes in the TMGSig-high subtype. (B) Ocoprint representation of the top 20 most frequently mutated genes in the TMGSig-low subtype.

10^{-13}). The C1 cluster comprised a substantial proportion of the high-risk group. The intersection of these two classifications may be useful for identifying outpatients with poor prognoses.

Our KEGG analysis of 11 TM-DEGs showed that they were particularly significant in the FoxO signaling pathway, progesterone-mediated oocyte maturation, and oocyte meiosis. In ovarian and age-related infertility, one of the factors that determine the success of IVF is telomeres. Fer-

tilization will not succeed if the oocyte's telomere length is shortened [49]. Pharmacological treatment of short telomeres has been used in various aging and degenerative diseases [50,51], and follicular cells can be considered potential targets for treatments to improve aging oocytes by promoting increases in telomere lengths. Cancer, as an age-related disease, may also serve as a reference to the application of telomere therapy for aging diseases.

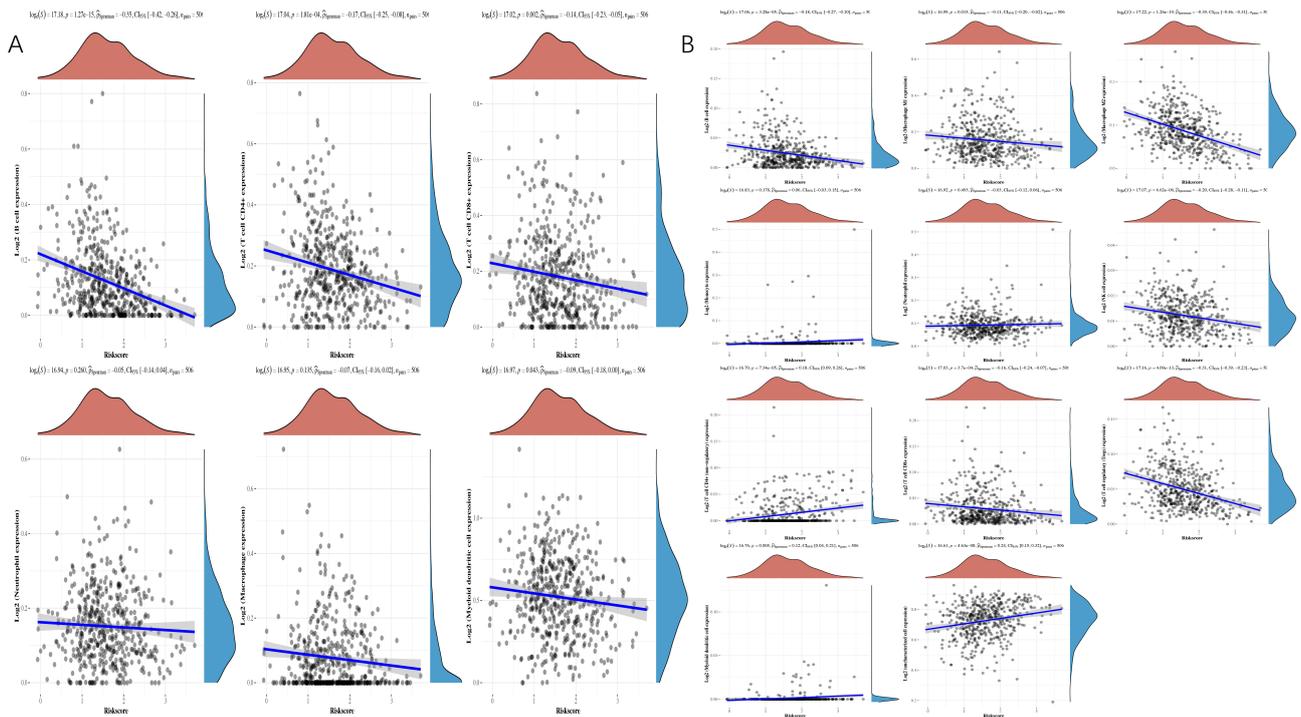


Fig. 12. Spearman correlation analysis between risk score and immune score. (A) TIMER algorithm. (B) quanTiseq algorithm.

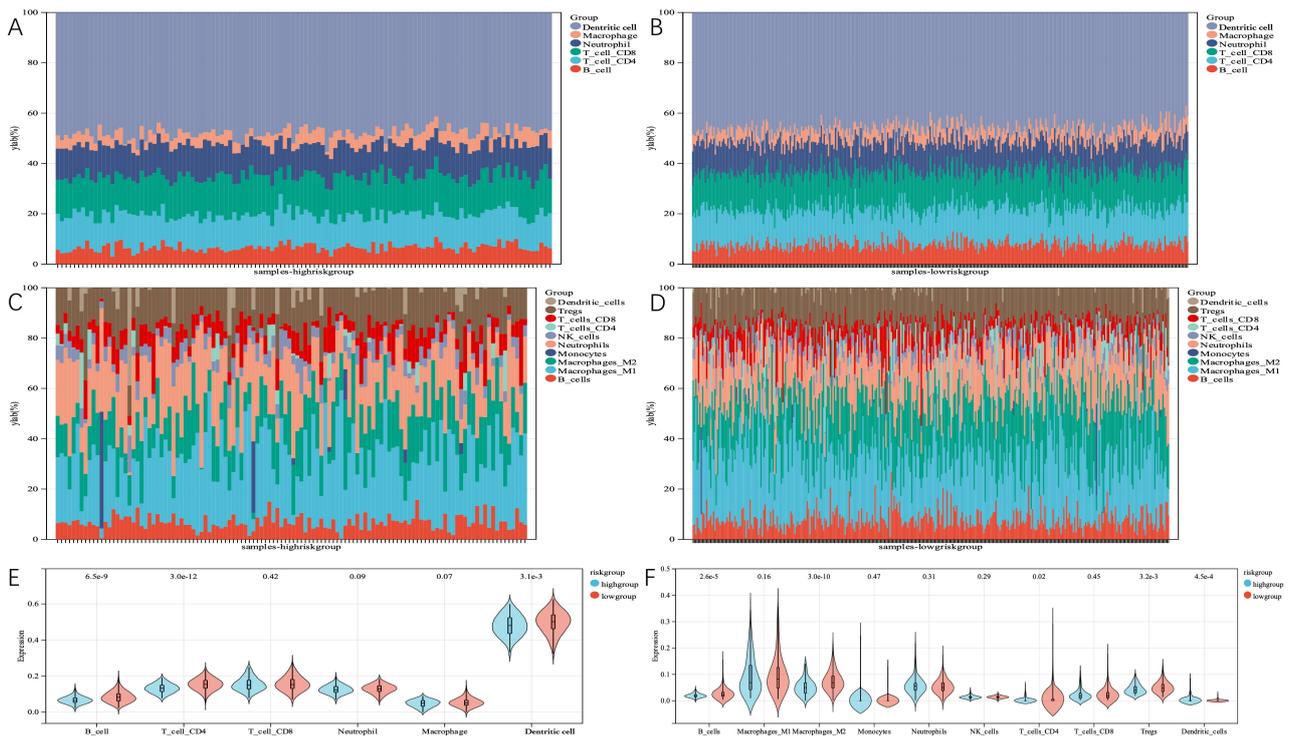


Fig. 13. The immune landscape of TMGsig-high and TMGsig-low subtypes. (A) The relative abundance of 6 cell types in the high-risk score subgroup using the TIMER algorithm. (B) The relative abundance of 6 cell types in the low-risk score subgroup using the TIMER algorithm. (C) The relative abundance of 10 cell types in the high-risk score subgroups was calculated by the quanTiseq algorithm. (D) The relative abundance of 10 cell types in the low-risk score subgroups was calculated by the quanTiseq algorithm. (E) Boxplot was used to compare the differences in cell infiltration abundance between the two subgroups using the TIMER algorithm. (F) Boxplot was used to compare the differences in cell infiltration abundance between the two subgroups using the quanTiseq algorithm.

With the rapid development of tumor immunotherapy, lung cancer therapy has entered a new era of treatment. Based on the correlation between lung cancer, immunity, and telomeres, we continued to analyze the relationship between TMGs and TMEs. We found that risk scores were generally negatively correlated with immune cells, the most significant of which were B cells, macrophage M2 cells, and regulatory T cells. A team of researchers led by Alessio Lanna of University College London recently published an interesting study in *Nature Cell Biology*, detailing the mechanism whereby memory T cells extend their lifespans. They found that, during antigen presentation, antigen-presenting cells (APCs) actively “donate” their telomeres to T cells via vesicles, to help them prolong their lifespans [52]. Combined with the results of this study, the increase in risk score and the decrease in APCs, composed of B cells, dendritic cells, or macrophages, lead to the shortening of T cell lifespans, the decline of immune function, and worse prognoses for lung cancer patients.

5. Conclusions

Using the TCGA dataset, we created a telomere-related gene risk signature and validated it using three GEO datasets. The results showed that this signature was very good at distinguishing patients’ prognoses. Our findings demonstrate a correlation between TMG subtypes in LUAD and alterations in the immune TME. These findings may be advantageous for immunotherapy interventions in patients with LUAD.

Availability of Data and Materials

The datasets presented in this study can be found in online repositories and additional inquiries can be directed to the corresponding author upon reasonable request.

Author Contributions

First draft writing: XL and JL; Methodology: JW; Data analysis: XL and JL; Review and editing: XG and MD; Conception design and review: SS and JW; Data acquisition and assisted analysis: XG and MD. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

Not applicable.

Funding

This research was funded by Jiangyin Science and Technology Bureau Social Development Science and Technology Demonstration Project, grant number JY0603A021014210005PB.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a Cancer Journal for Clinicians*. 2021; 71: 209–249.
- [2] Ma J, Ward EM, Smith R, Jemal A. Annual number of lung cancer deaths potentially avertable by screening in the United States. *Cancer*. 2013; 119: 1381–1385.
- [3] Ye Y, Dai Q, Li S, He J, Qi H. A Novel Defined Risk Signature of the Ferroptosis-Related Genes for Predicting the Prognosis of Ovarian Cancer. *Frontiers in Molecular Biosciences*. 2021; 8: 645845.
- [4] Perry PJ, Arnold JR, Jenkins TC. Telomerase inhibitors for the treatment of cancer: the current perspective. *Expert Opinion on Investigational Drugs*. 2001; 10: 2141–2156.
- [5] Sugarman ET, Zhang G, Shay JW. In perspective: an update on telomere targeting in cancer. *Molecular Carcinogenesis*. 2019; 58: 1581–1588.
- [6] Jang JS, Choi YY, Lee WK, Choi JE, Cha SI, Kim YJ, *et al.* Telomere length and the risk of lung cancer. *Cancer Science*. 2008; 99: 1385–1389.
- [7] Doherty JA, Grieshaber L, Houck JR, Barnett MJ, De Dieu Tapsoba J, Thornquist MD, *et al.* Nested case-control study of telomere length and lung cancer risk among heavy smokers in the β -Carotene and Retinol Efficacy Trial. *British Journal of Cancer*. 2018; 118: 1513–1517.
- [8] Yuan J, Beckman KB, Wang R, Bull C, Adams-Haduch J, Huang JY, *et al.* Leukocyte telomere length in relation to risk of lung adenocarcinoma incidence: Findings from the Singapore Chinese Health Study. *International Journal of Cancer*. 2018; 142: 2234–2243.
- [9] Weischer M, Nordestgaard BG, Cawthon RM, Freiberg JJ, Tybjaerg-Hansen A, Bojesen SE. Short Telomere Length, Cancer Survival, and Cancer Risk in 47102 Individuals. *Journal of the National Cancer Institute*. 2013; 105: 459–468.
- [10] Mons U, Müezziner A, Schöttker B, Dieffenbach AK, Butterbach K, Schick M, *et al.* Leukocyte Telomere Length and all-Cause, Cardiovascular Disease, and Cancer Mortality: Results from Individual-Participant-Data Meta-Analysis of 2 Large Prospective Cohort Studies. *American Journal of Epidemiology*. 2017; 185: 1317–1326.
- [11] Needham BL, Rehkopf D, Adler N, Gregorich S, Lin J, Blackburn EH, *et al.* Leukocyte telomere length and mortality in the National Health and Nutrition Examination Survey, 1999–2002. *Epidemiology*. 2015; 26: 528–535.
- [12] Shen W, Song Z, Zhong X, Huang M, Shen D, Gao P, *et al.* Sangerbox: a comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta*. 2022; 1: e36.
- [13] Braun DM, Chung I, Kepper N, Deeg KI, Rippe K. TelNet - a database for human and yeast genes involved in telomere maintenance. *BMC Genetics*. 2018; 19: 32.
- [14] Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010; 26: 1572–1573.
- [15] Siegel RL, Miller KD, Fuchs HE, Jemal A. *Cancer Statistics, 2021*. *CA: a Cancer Journal for Clinicians*. 2021; 71: 7–33.
- [16] Rux DR, Wellik DM. Hox genes in the adult skeleton: Novel functions beyond embryonic development. *Developmental Dynamics*. 2017; 246: 310–317.
- [17] Dunwell TL, Holland PW. Diversity of human and mouse homeobox gene expression in development and adult tissues. *BMC Developmental Biology*. 2016; 16: 40.

- [18] Abe M, Hamada J, Takahashi O, Takahashi Y, Tada M, Miyamoto M, *et al.* Disordered expression of HOX genes in human non-small cell lung cancer. *Oncology Reports*. 2006; 15: 797–802.
- [19] Tang B, Qi G, Sun X, Tang F, Yuan S, Wang Z, *et al.* HOXA7 plays a critical role in metastasis of liver cancer associated with activation of Snail. *Molecular Cancer*. 2016; 15: 57.
- [20] Lee S, Jang C, Lee K. Polo-Like Kinases (Plks), a Key Regulator of Cell Cycle and New Potential Target for Cancer Therapy. *Development and Reproduction*. 2014; 18: 65–71.
- [21] Zou Y, Zhang Q, Zhang J, Chen X, Zhou W, Yang Z, *et al.* A common indel polymorphism of the Desmoglein-2 (DSG2) is associated with sudden cardiac death in Chinese populations. *Forensic Science International*. 2019; 301: 382–387.
- [22] Fu BM. Tumor Metastasis in the Microcirculation. *Advances in Experimental Medicine and Biology*. 2018; 1097: 201–218.
- [23] Cai F, Zhu Q, Miao Y, Shen S, Su X, Shi Y. Desmoglein-2 is overexpressed in non-small cell lung cancer tissues and its knockdown suppresses NSCLC growth by regulation of p27 and CDK2. *Journal of Cancer Research and Clinical Oncology*. 2017; 143: 59–69.
- [24] Li G, Yin H, Kuret J. Casein Kinase 1 Delta Phosphorylates Tau and Disrupts its Binding to Microtubules. *Journal of Biological Chemistry*. 2004; 279: 15938–15945.
- [25] Cheong JK, Hung NT, Wang H, Tan P, Voorhoeve PM, Lee SH, *et al.* IC261 induces cell cycle arrest and apoptosis of human cancer cells via CK1 δ/ϵ and Wnt/ β -catenin independent inhibition of mitotic spindle formation. *Oncogene*. 2011; 30: 2558–2569.
- [26] Tiong K, Chang K, Yeh K, Liu T, Wu J, Hsieh P, *et al.* CSNK1E/CTNBN1 are Synthetic Lethal to TP53 in Colorectal Cancer and are Markers for Prognosis. *Neoplasia*. 2014; 16: 441–450.
- [27] Grigoryan T, Wend P, Klaus A, Birchmeier W. Deciphering the function of canonical Wnt signals in development and disease: conditional loss- and gain-of-function mutations of beta-catenin in mice. *Genes & Development*. 2008; 22:2308–2341.
- [28] Rivlin N, Brosh R, Oren M, Rotter V. Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. *Genes & Cancer*. 2011; 2: 466–474.
- [29] Khan MT, Prajapati B, Lakhina S, Sharma M, Prajapati S, Chosdol K, *et al.* Identification of Gender-Specific Molecular Differences in Glioblastoma (GBM) and Low-Grade Glioma (LGG) by the Analysis of Large Transcriptomic and Epigenomic Datasets. *Frontiers in Oncology*. 2021; 11: 699594.
- [30] Jiang Z, Wang X, Li J, Yang H, Lin X. Aldolase a as a prognostic factor and mediator of progression via inducing epithelial-mesenchymal transition in gastric cancer. *Journal of Cellular and Molecular Medicine*. 2018; 22: 4377–4386.
- [31] Tian W, Zhou J, Chen M, Qiu L, Li Y, Zhang W, *et al.* Bioinformatics analysis of the role of aldolase A in tumor prognosis and immunity. *Scientific Reports*. 2022; 12: 11632.
- [32] He Y, Deng F, Zhao S, Zhong S, Zhao J, Wang D, *et al.* Analysis of miRNA-mRNA network reveals miR-140-5p as a suppressor of breast cancer glycolysis via targeting GLUT1. *Epigenomics*. 2019; 11: 1021–1036.
- [33] Zhao C, Zhou Y, Ma H, Wang J, Guo H, Liu H. A four-hypoxia-genes-based prognostic signature for oral squamous cell carcinoma. *BMC Oral Health*. 2021; 21: 232.
- [34] Kawai K, Uemura M, Munakata K, Takahashi H, Haraguchi N, Nishimura J, *et al.* Fructose-bisphosphate aldolase a is a key regulator of hypoxic adaptation in colorectal cancer cells and involved in treatment resistance and poor prognosis. *International Journal of Oncology*. 2017; 50: 525–534.
- [35] Tang Y, Yang X, Feng K, Hu C, Li S. High expression of aldolase a is associated with tumor progression and poor prognosis in hepatocellular carcinoma. *Journal of Gastrointestinal Oncology*. 2021; 12: 174–183.
- [36] Chang Y, Chan Y, Chang W, Lin Y, Yang C, Su C, *et al.* Feedback regulation of ALDOA activates the HIF-1 α MMP9 axis to promote lung cancer progression. *Cancer Letters*. 2017; 403: 28–36.
- [37] Fu H, Gao H, Qi X, Zhao L, Wu D, Bai Y, *et al.* Aldolase A promotes proliferation and G1/S transition via the EGFR/MAPK pathway in non-small cell lung cancer. *Cancer Communications*. 2018; 38: 18.
- [38] Chang YC, Yang YF, Chiou J, Tsai HF, Fang CY, Yang CJ, *et al.* Nonenzymatic function of Aldolase A downregulates miR-145 to promote the Oct4/DUSP4/TRAF4 axis and the acquisition of lung cancer stemness. *Cell Death & Disease*. 2020; 11: 195.
- [39] Wang J, Sheng Z, Dong Z, Wu Q, Cai Y. The mechanism of radiotherapy for lung adenocarcinoma in promoting protein SIRT6-mediated deacetylation of RBBP8 to enhance the sensitivity of targeted therapy. *International Journal of Immunopathology and Pharmacology*. 2022; 36: 039463202211307.
- [40] Hong Z, Jiang J, Lan L, Nakajima S, Kanno S, Koseki H, *et al.* A polycomb group protein, PHF1, is involved in the response to DNA double-strand breaks in human cell. *Nucleic Acids Research*. 2008; 36: 2939–2947.
- [41] Brien GL, Healy E, Jerman E, Conway E, Fadda E, O'Donovan D, *et al.* A chromatin-independent role of Polycomb-like 1 to stabilize p53 and promote cellular quiescence. *Genes and Development*. 2015; 29: 2231–2243.
- [42] Ning J, Wang F, Bu J, Zhu K, Liu W. Down-regulated m6A reader FTO destabilizes PHF1 that triggers enhanced stemness capacity and tumor progression in lung adenocarcinoma. *Cell Death Discovery*. 2022; 8: 354.
- [43] Mazzei F, Viel A, Bignami M. Role of MUTYH in human cancer. *Mutation Research*. 2013; 743–744: 33–43.
- [44] Agustina L, Hahm S, Han SH, Tran AHV, Chung JH, Park J, *et al.* Visualization of the physical and functional interaction between hMYH and hRad9 by Dronpa bimolecular fluorescence complementation. *BMC Molecular Biology*. 2014; 15: 17.
- [45] Zhu M, Chen X, Zhang H, Xiao N, Zhu C, He Q, *et al.* AluYb8 insertion in the MUTYH gene and risk of early-onset breast and gastric cancers in the Chinese population. *Asian Pacific Journal of Cancer Prevention*. 2011; 12: 1451–1455.
- [46] D'Agostino VG, Minoprio A, Torrerri P, Marinoni I, Bossa C, Petrucci TC, *et al.* Functional analysis of MUTYH mutated proteins associated with familial adenomatous polyposis. *DNA Repair*. 2010; 9: 700–707.
- [47] Raetz AG, Xie Y, Kundu S, Brinkmeyer MK, Chang C, David SS. Cancer-associated variants and a common polymorphism of MUTYH exhibit reduced repair of oxidative DNA damage using a GFP-based assay in mammalian cells. *Carcinogenesis*. 2012; 33: 2301–2309.
- [48] Ungar L, Yosef N, Sela Y, Sharan R, Ruppin E, Kupiec M. A genome-wide screen for essential yeast genes that affect telomere length maintenance. *Nucleic Acids Research*. 2009; 37: 3840–3849.
- [49] Liu L, Trimarchi JR, Smith PJS, Keefe DL. Checkpoint for DNA integrity at the first mitosis after oocyte activation. *Molecular Reproduction and Development*. 2002; 62: 277–288.
- [50] Tsoukalas D, Buga AM, Docea AO, Sarandi E, Mitrut R, Renieri E, *et al.* Reversal of brain aging by targeting telomerase: A nutraceutical approach. *International Journal of Molecular Medicine*. 2021; 48: 199.
- [51] Prieto-Oliveira P. Telomerase activation in the treatment of aging or degenerative diseases: a systematic review. *Molecular and Cellular Biochemistry*. 2021; 476: 599–607.
- [52] Lanna A, Vaz B, D'Amra C, Valvo S, Vuotto C, Chirchiù V, *et al.* An intercellular transfer of telomeres rescues T cells from senescence and promotes long-term immunological memory. *Nature Cell Biology*. 2022; 24: 1461–1474.