Research Article

Identification and Analysis of Senescence-Related Genes in Head and Neck Squamous Cell Carcinoma by a Comprehensive Bioinformatics Approach

Lin Deng [iD],[1,2] **Jinglin Mi,**[1] **Xiaolan Ruan,**[2] **Guozhen Zhang,**[2] **Yufei Pan** [iD],[2] **and Rensheng Wang** [iD][1]

[1]*Department of Oncology, The First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China*
[2]*Department of Radiation Oncology, Nanxishan Hospital of Guangxi Zhuang Autonomous Region, Guilin 541004, China*

Correspondence should be addressed to Yufei Pan; 23865934@qq.com and Rensheng Wang; 13807806008@163.com

Lin Deng and Jinglin Mi contributed equally to this work.

Head and neck cancer is the sixth most frequent cancer all over the world, with the majority of subtypes of head and neck squamous cell carcinoma (HNSCC). Cellular senescence-associated genes have been confirmed to play a critical role in cancer and have the potential to be prognostic biomarkers for cancer. Clinical information of HNSCC samples and expression data were acquired from public databases. Expression profiles of genes related to cellular senescence were used to identify molecular subtypes by consensus clustering. To screen differentially expressed genes (DEGs) between different subtypes, differential analysis was performed. We used the univariate Cox regression to identify prognostic DEGs and performed least absolute shrinkage and selection operator (LASSO) to optimize and construct a prognostic model. CIBERSORT, ESTIMATE, and TIDE tools were applied to estimate immune characteristics. Four molecular subtypes were established based on cellular senescence-associated genes. Differential prognosis was observed among different subtypes with C4 having the longest overall survival and C1 having the worst prognosis. C4 subtype also showed the highest immune infiltration. We screened a total of eight cellular senescence prognosis-related genes and established a cellular senescence-related signature score (CSRS.Score) that could stratify samples into high-CSRS.Score and low-CSRS.Score groups. The high-CSRS.Score group had worse prognosis, lower immune infiltration, and lower response to immunotherapy. We further improved the prognostic model and survival prediction by combining CSRS.Score with clinicopathological features using a decision tree model, which had high predictive accuracy and survival prediction. This study demonstrated an important role of cellular senescence in HNSCC. The identified eight cellular senescence-associated genes have the potential to provide ideas for adjuvant treatment and personalized treatment of HNSCC patients.

1. Introduction

Head and neck cancer (HNC) is the sixth most frequently diagnosed cancer type that causes 500,000 affected individuals per year worldwide [1]. Head and neck squamous cell carcinoma (HNSCC) accounts for the majority HNC patients, and more than half of the patients with HNSCC are initially diagnosed with locally advanced disease [2, 3]. Lymph node (LN) metastasis is a negative signal of head and neck cancer prognosis. However, it is challengeable to identify metastatic LN within the fibroadipose tissue [4]. The prognosis for HNSCC remains poor even with the use of combination therapy including surgery, radiation, chemotherapy, and immunotherapy [5]. Although no tumor LN is detected from clinical and radiographic estimation, there is still a high possibility over than 30% to observe nodal metastasis in the surgery

[6]. Therefore, there is an urgent need to provide effective biomarkers for early diagnosis, personalized treatment, and prognostic evaluation.

Senescence is a nearly unavoidable feature in all creatures, which is marked by a descending function of multiple cells and tissues. In spite of that degeneration is the most common age-related phenotype, aging allows to generate gain-of-function changes that lead to abnormal cell proliferation [7]. Moreover, these changes can result in genomic instability that enable to provide an advantage for the abnormal cells in proliferation, migration, and escape from immune surveillance [8]. Obviously, these phenotypes are the hallmarks of malignant cancers. Senescence plays a two-sided role in cancer development, preventing tumorigenesis by cell growth arrest in precancerous cells, but also facilitating malignant transformation of adjacent cells through protumorigenic drivers [9, 10]. Senescent cells can alter epigenetic modifications in neighboring cells by releasing senescence-associated signals [11–13]. A number of genes have been demonstrated to regulate senescence in cancer cells, such as p53 [14], Raf1 [14], MAP2K6/p38 [15], and PTEN [16]. Therapy-induced senescence has been observed in cancer cells after radiotherapy or chemotherapy [17]. When exposed to various conventional and targeted anticancer drugs, tumor cell senescence is induced, resulting in a positive effect on patient treatment [18, 19]. Thus, senescence is considered as a therapeutic target for clinical cancer treatment [20].

Senescence-associated genes also have the great potential to predict cancer prognosis. Althubiti et al. identified 10 plasma membrane-associated proteins expressed in senescent cells to be prognostic biomarkers especially in breast cancer [21]. Coppola et al. discovered a series of senescence-associated genes that correlated with age, overall survival, and grade of glioblastoma [22]. Yang et al. identified seven age-related genes by analyzing the expression profiles of HNSCC and adjacent cancer samples [23]. The risk score based on the seven age-related genes was significantly related to prognosis and immune response. But none of the studies have explored a molecular subtyping system based on senescence-associated genes in HNSCC. Senescence-associated molecular subtypes may help to further understand the role of cellular senescence in the tumor progression.

Therefore, in this study, we used cellular senescence-associated genes to identify molecular subtypes. Differential pathways and immune features were observed among different subtypes. Differential expressed genes (DEGs) were screened between different subtypes, and least absolute shrinkage and selection operator (LASSO) regression analysis was used to develop a cellular senescence-related signature scoring (CSRS) system. The CSRS system could define CSRS.Score for each HNSCC sample and classify them into high-CSRS.Score and low-CSRS.Score groups. Importantly, CSRS.Score had the potential to guide immunotherapy and chemotherapy for HNSCC patients. A decision tree and a nomogram based on CSRS.Score were constructed to more accurately predict prognosis than CSRS.Score only.

2. Materials and Methods

2.1. Data Source and Preprocessing. From The Cancer Genome Atlas (TCGA) database (named as TCGA cohort), we downloaded RNA-seq data of HNSCC samples and removed samples that did not have survival time, clinical follow-up information, or status of patients' survival. Ensembl ID was transformed into gene symbol. The median value of gene expression was selected for the genes with multiple gene symbols. GSE65858 and GSE41613 cohorts including gene expression profiles of HNSCC samples were obtained from Gene Expression Omnibus (GEO) database and were used as validation cohorts. We downloaded the annotation information of the corresponding microarray platform and mapped the probes to genes based on the annotation information to remove the probes that match one probe to multiple genes. If certain number probes matched to one gene, the median value was taken as the expression value of that gene. Finally, 499, 253, and 97 samples were remained in TCGA, GSE65858, and GSE41613 cohorts, respectively.

From CellAge database (https://genomics.senescence .info/cells/), we obtained 279 cellular senescence-associated genes.

2.2. Molecular Typing Based on Cellular Senescence-Associated Genes. We next constructed a consistency matrix by ConsensusClusterPlus to cluster HNSCC samples [24]. The expression data of genes associated with cellular senescence was used as a basis to obtain the molecular subtypes of the samples. "Pam" algorithm and "1 - Spearman correlation" were determined as a metric distance to perform 500 bootstraps, with each bootstrap consisting of 80% patients in the training set. Cluster numbers were set from 2 to 10, and the optimal cluster was determined by cumulative distribution function (CDF) and consensus matrix. Finally, the confirmed clusters were the molecular subtypes.

2.3. Construction of a CSRS.Score Scoring System. We identified differentially expressed cellular senescence genes between subtypes using limma R package [25] and selected prognostically significant differentially expressed genes through the univariate Cox regression analysis ($P < 0.05$). LASSO regression using the glmnet R package [26] and stepwise Akaike information criterion (stepAIC) [27] were performed to compress and reduce the differential genes to obtain prognostic genes associated with cellular senescence. The CSRS.Score for each sample was shown as follows: CSRS.Score = $\Sigma \beta i \times$ Expi, where Expi indicates the gene expression level of prognostic genes and β i is Cox regression coefficients of the corresponding genes. CSRS.Score was normalized using z-score, and the threshold "0" was determined to classify samples into low-risk and high-risk groups. The Kaplan-Meier (KM) survival analysis was conducted to assess the overall survival of different molecular subtypes. Significant differences were determined using the log-rank test.

2.4. Assessment of Immune Infiltration. CIBERSORT algorithm was employed to estimate the proportion of 22 immune cell types [28]. ESTIMATE algorithm was used to

calculate stromal score and immune score for evaluating stromal and immune infiltration [29].

2.5. Prediction of Immunotherapy Responsiveness.

We used the TIDE algorithm to validate predicted treatment responsiveness. The TIDE algorithm is a computational method for predicting immune checkpoint blockade responsiveness using gene expression profiles [30]. TIDE can use gene expression information to predict cancer sensitivity to immune checkpoint therapy. The TIDE algorithm evaluates three immunosuppressive cell types that limit T cell infiltration in tumors, including M2 tumor-associated macrophages (TAMs), tumor-associated fibroblasts (CAFs), and myeloid-derived suppressor cells (MDSCs). The dysfunction score of tumor-infiltrating cytotoxic T lymphocytes (CTLs) (T cell dysfunction) and the exclusion score of CTLs by immunosuppressive factors (T cell exclusion) can be calculated by TIDE analysis.

2.6. Gene Set Enrichment Analysis (GSEA).

GSEA allows to calculate the enrichment score of a gene set for annotating biological function [31]. We used GSEA for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis in different molecular subtypes. Gene sets of KEGG pathways were accessed from the Molecular Signature Database (MSigDB) [32]. The enrichment scores of aneuploidy, homologous recombination defect number of segments, and fraction altered were calculated using GSEA.

2.7. Statistical Analysis.

R software (v4.1) was applied to conduct all statistical analysis. The Kruskal-Wallis test was performed in testing the significance among four subtypes. Between high- and low-risk groups, the Wilcoxon test was performed to test the significance. Log-rank test was conducted in the Cox regression analysis and survival analysis. ANOVA was conducted in comparing different groups containing multiple subgroups. $P < 0.05$ was considered as significant.

3. Results

3.1. Molecular Typing Based on Cellular Senescence-Associated Genes.

First, we extracted the expression of cellular senescence-associated genes from TCGA cohort. Then, the univariate Cox regression analysis was performed, and we obtained 28 genes associated with prognosis (Table S1) ($P < 0.01$). Then, based on the expression data of 28 prognosis-related cellular senescence-associated genes, we clustered 499 HNSCC samples into four clusters (molecular subtypes) through the determination of the CDF and the CDF delta area (Figures 1(a) and 1(b)). The consensus matrix showed that four clusters were independently distributed for most samples (Figure 1(c)). KM survival curves displayed that four molecular subtypes had significant differences of overall survival ($P < 0.01$, Figure 1(d)), with C4 having the most favorable prognosis and C1 having the highest proportion of dead samples ($P < 0.05$, Figure 1(e)). In addition, C1 also had a higher proportion of advanced stages of T stage, N stage, and AJCC stage (Figure S1).

3.2. Genomic Characteristics and Enriched Pathways of Molecular Subtypes.

We explored the differences of genomic alterations in the TCGA cohort (acquired from previous research, [33]) among the four molecular subtypes. C1 subtype showed higher score of aneuploidy, homologous recombination defects, number of segments, and fraction altered (Figure 2(a)). However, no significant difference was shown in tumor mutation burden. In addition, we also analyzed the frequency of gene mutations among molecular subtypes (Figure 2(b)). TP53 had the highest mutation frequency of 81.3% and over a half samples had TP53 mutation. Missense mutation contributed the majority of gene mutation, while nonsense mutation was the most in CDKN2A.

Next, we analyzed whether differential enriched pathways exist in the different molecular subtypes by GSEA. 37 pathways were identified to be significantly enriched in the C1 subtype in the TCGA cohort. The enriched pathways mainly include cancer-related pathways, such as small cell lung cancer, ECM receptor interaction, and Wnt signaling pathway (Figure S2). The results suggested that cellular senescence-associated genes were possibly involved in cancer-related pathways and inflammatory pathways.

3.3. Immunological Characteristics and Immunotherapy/Chemotherapy Responses in Different Molecular Subtypes.

We used immune cell signatures to assess immune cell infiltration in different subtypes to evaluate their immune characteristics. CIBERSORT revealed that 16 of 22 immune cells had a significant difference among four subtypes, such as regulatory T cells, CD8 T cells, resting memory CD4 T cells, and M0 macrophages ($P < 0.05$, Figure 3(a)). C1 had the lowest stromal score and immune score, while C4 had the highest scores ($P < 0.0001$, Figure 3(b)), indicating higher immune infiltration in C4. We considered that differential immune characteristics of four subtypes may result in different immune responses to immunotherapy.

Therefore, we assessed the expression of immune checkpoint genes in different molecular subtypes. We could see that the majority immune checkpoints were differentially expressed among four subtypes (Figure 3(c)), suggesting that the different subtypes may differentially response to immune checkpoint blockade. Not surprisingly, TIDE analysis revealed different responses of four subtypes to immune checkpoint inhibitors. As shown in Figure 3(d), the highest TIDE score was shown in C1, indicating that C1 was more probably to escape from immunotherapy. A higher proportion of MDSCs and CAFs may result in a higher T cell exclusion and unsatisfied immune response (Figure 3(d)). In addition, we also analyzed the predicted response of different molecular subtypes to four chemotherapeutic drugs (paclitaxel, docetaxel, cisplatin, and 5-fluorouracil). C1 and C2 subtypes are more sensitive to paclitaxel, docetaxel, and cisplatin drugs (Figure 3(e)).

3.4. Identification of Key Cellular Senescence-Associated Genes.

We then used limma package to screen differentially expressed cellular senescence-associated genes between C1 and non-C1, C2 and non-C2, C3 and non-C3, and C4 and non-C4 molecular subtypes based on the conditions of false
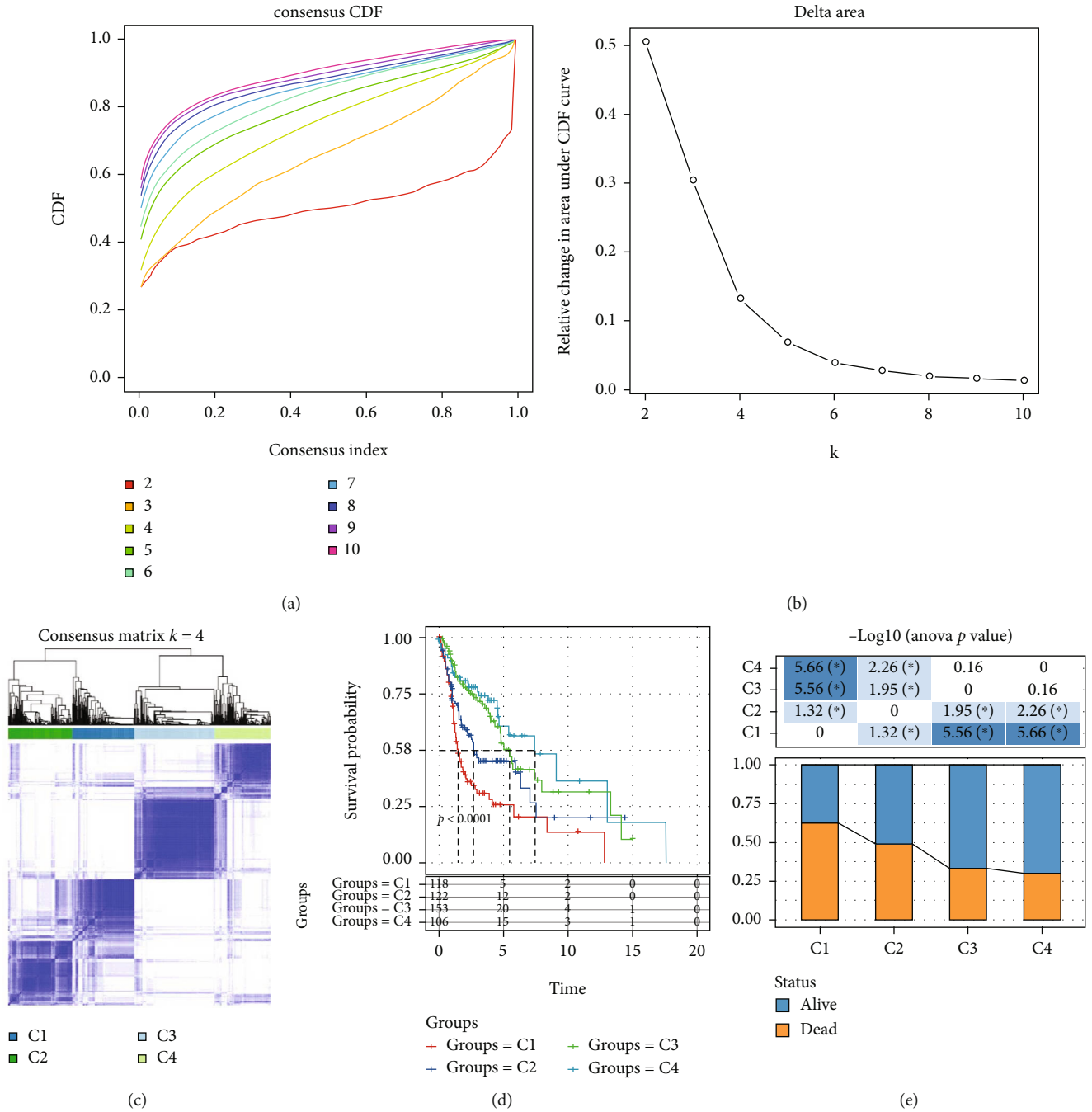
FIGURE 1: The TCGA cohort molecular typing based on cellular senescence-associated genes. (a) CDF curve for TCGA cohort samples. (b) CDF delta area curve for TCGA cohort samples. The vertical axis represents the relative change in area under CDF curve, and the horizontal axis represents the category number k. (c) At consensus $k = 4$, heat map of sample clustering. (d) KM curve of the four subtypes. (e) Survival status differences in different subtypes. $*P < 0.05$.

discovery rate (FDR) < 0.05 and $|\log 2(\text{fold change})| > 1$. 232 DEGs were identified by the above intersection. And 77 prognostic genes were confirmed by the univariate Cox regression analysis including 31 "risk" and 46 "protective" genes (Figure 4(a)). Next, LASSO regression was conducted to compress the 77 prognostic genes. As shown in Figure 4(b), the coefficients of the prognostic genes tended to zero as lambda increased, and the optimal model was confirmed when lambda $= 0.0275$ (Figure 4(c)). StepAIC

was further performed to optimize the model with the least number of prognostic genes. Finally, we identified eight key cellular senescence-associated genes related to prognosis (Figure 4(c)), including PYGL, KRT8, AREG, MAGEA4, DES, EPHX3, CDKN2A, and SPINK6.

3.5. Validation of the Eight-Gene Prognostic Model. We calculated and normalized the cellular senescence-related signature score (CSRS.Score) for each sample according to the
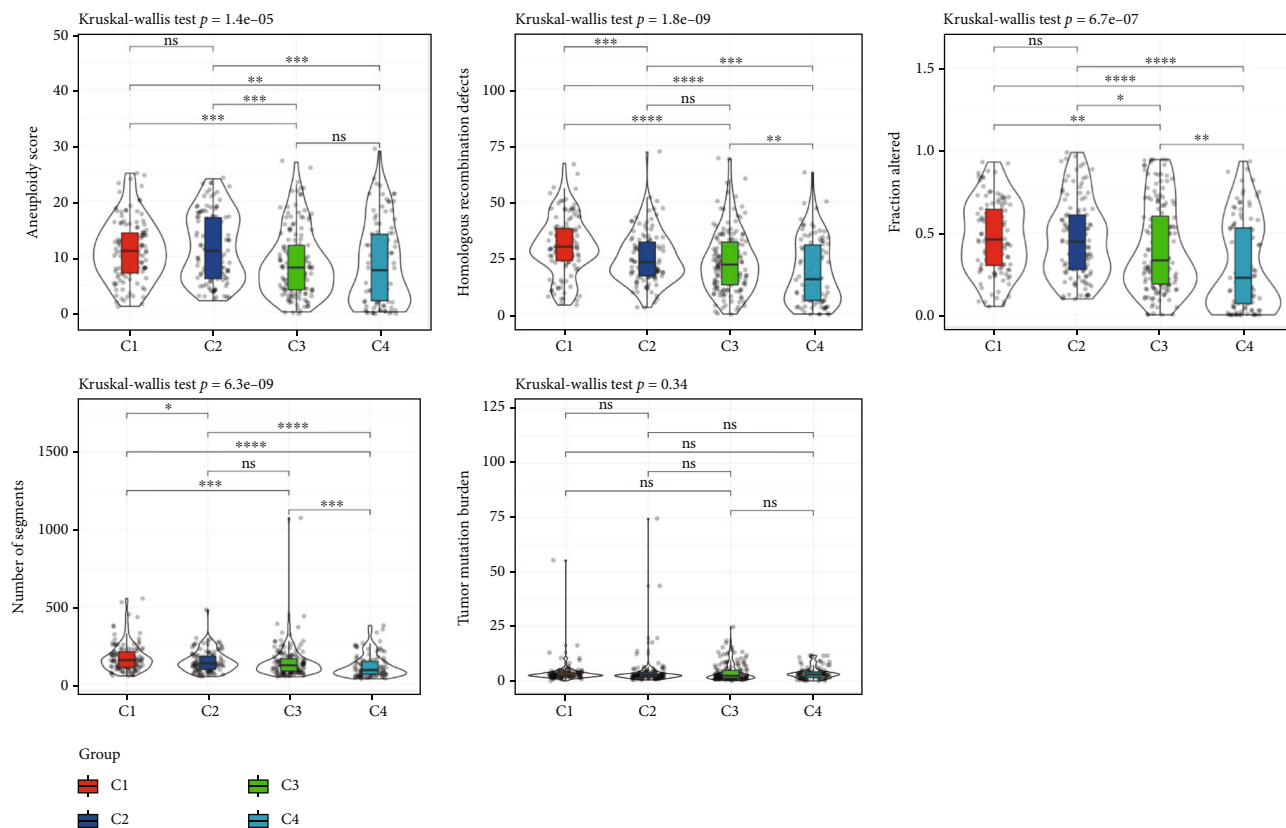
(a)
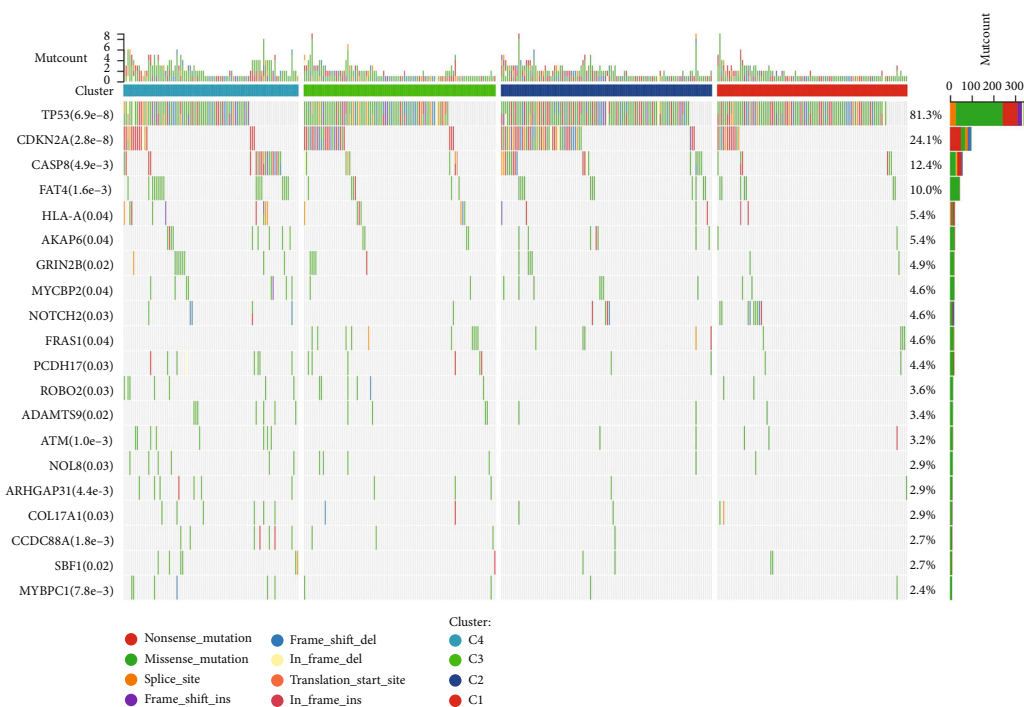
(b)

FIGURE 2: Genomic alterations in the molecular subtypes of the TCGA cohort. (a) Comparison of tumor mutation burden differences, aneuploidy score, homologous recombination defects, number of segments, and fraction altered. (b) Somatic mutations in the four molecular subtypes (chi-square test). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.
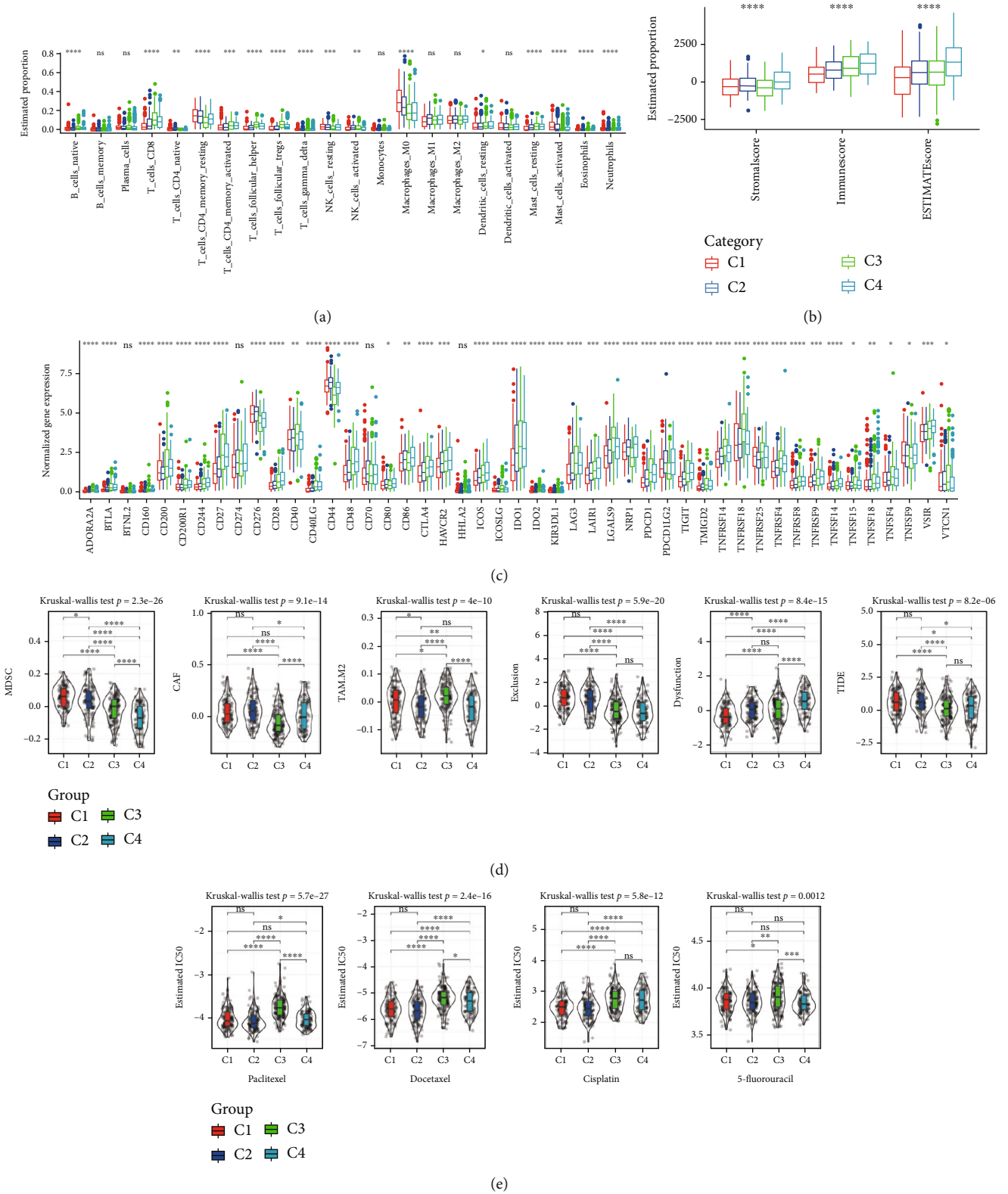
(a)

(b)

(c)

(d)

(e)

FIGURE 3: Immune characteristics and immunotherapy/chemotherapy differences between molecular subtypes. (a) The TCGA cohort differences in 22 immune cell scores between molecular subtypes. (b) Differences in ESTIMATE immune infiltration between molecular subtypes in the TCGA cohort. (c) Differentially expressed immune checkpoints between subtypes in the TCGA cohort. (d) Differences in TIDE analysis results between different subgroups in the TCGA cohort. (e) The box plots of the estimated IC50 for paclitaxel, docetaxel, cisplatin, and 5-fluorouracil in the TCGA cohort.
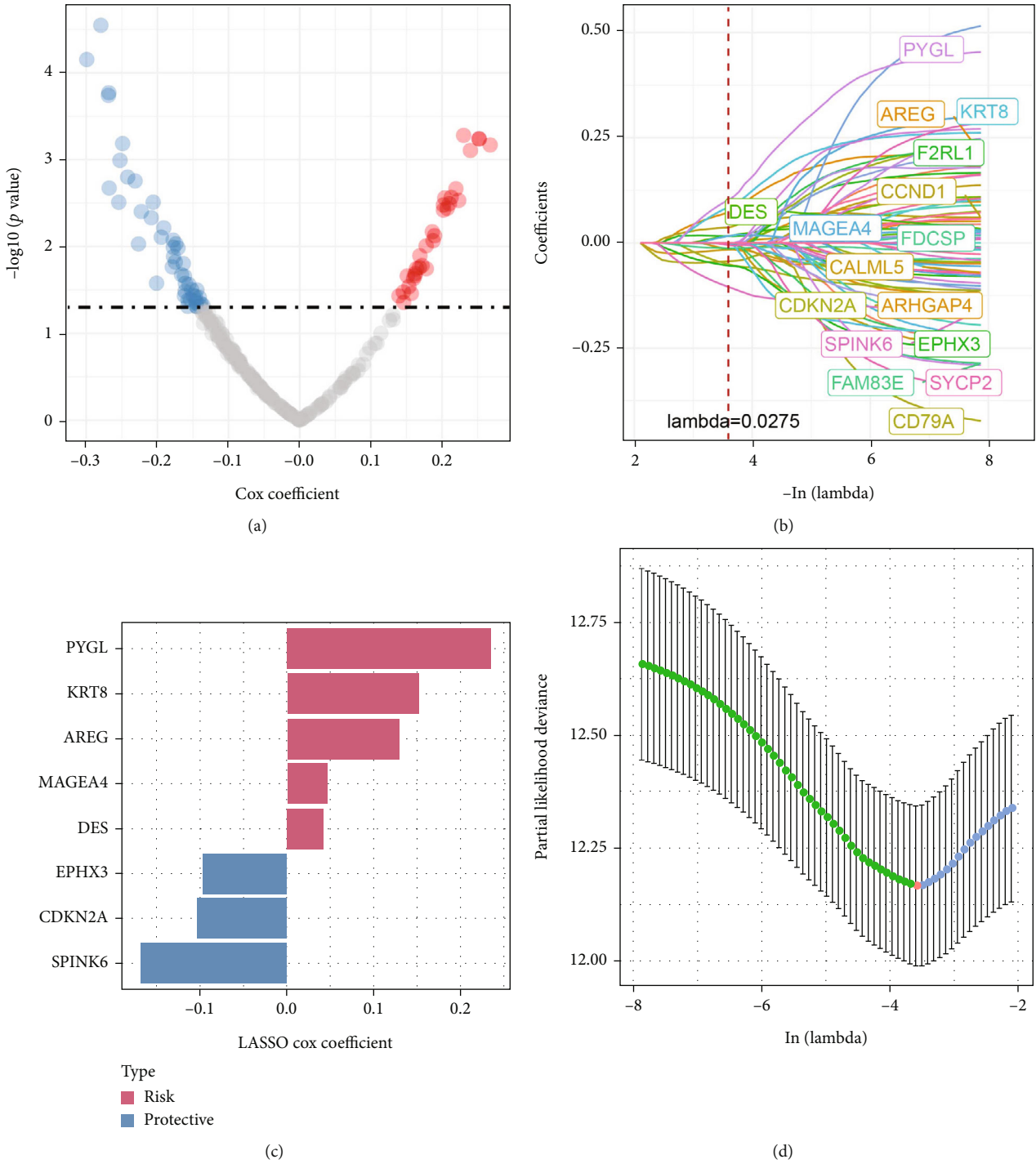
(a)

(b)

(c)

(d)

FIGURE 4: Identification of key cellular senescence-associated genes. (a) A total of 77 promising candidates were identified among the DEGs. (b) Trajectory of each independent variable with lambda. (c) Confidence interval under lambda. (d) LASSO coefficient distribution of gene signature correlated with the senescence.

eight-gene model. We classified the samples as the high-risk group if the score was greater than 0 and as the low-risk group if the score was less than 0. The distribution of CSRS.Score for samples in the training set (TCGA cohort) is shown in Figure 5(a). The high-risk group had an obviously higher proportion of dead samples and had a significantly short overall survival. We analyzed the prognostic predictive classification efficiency for 1 year (AUC = 0.72),

3 years (AUC = 0.72), and 5 years (AUC = 0.72), as shown in Figure 5(b). The KM survival curve showed that the overall survival of two risk groups was significantly different ($P < 0.0001$), and higher CSRS.Score had worse overall survival in the training cohort (Figure 5(c)). We further validated the robustness of the eight-gene model in two validation cohorts (GSE65858 and GSE41613), and the similar results were observed (Figures 5(d)–5(g)). For the
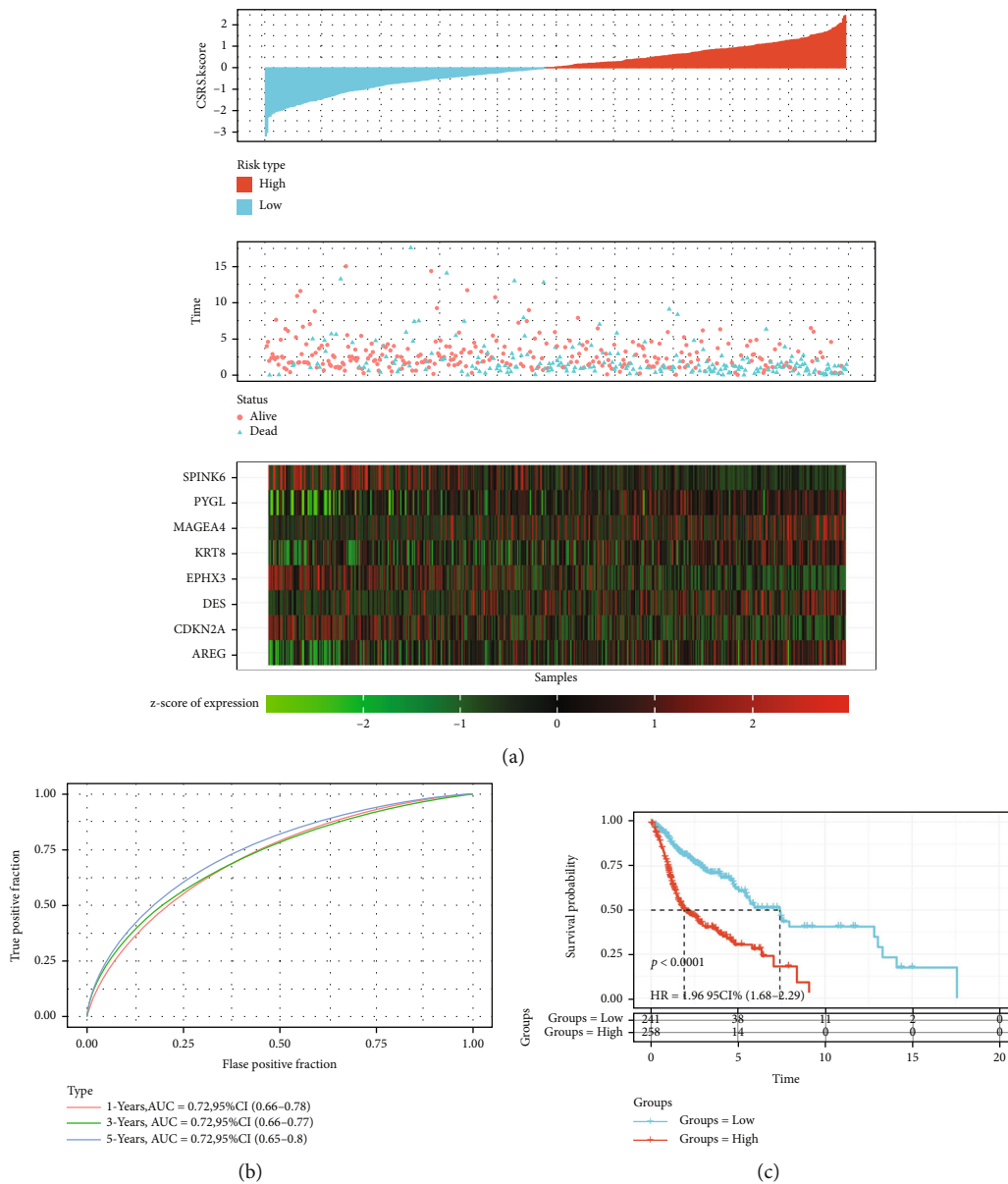
(a)

(b)

(c)
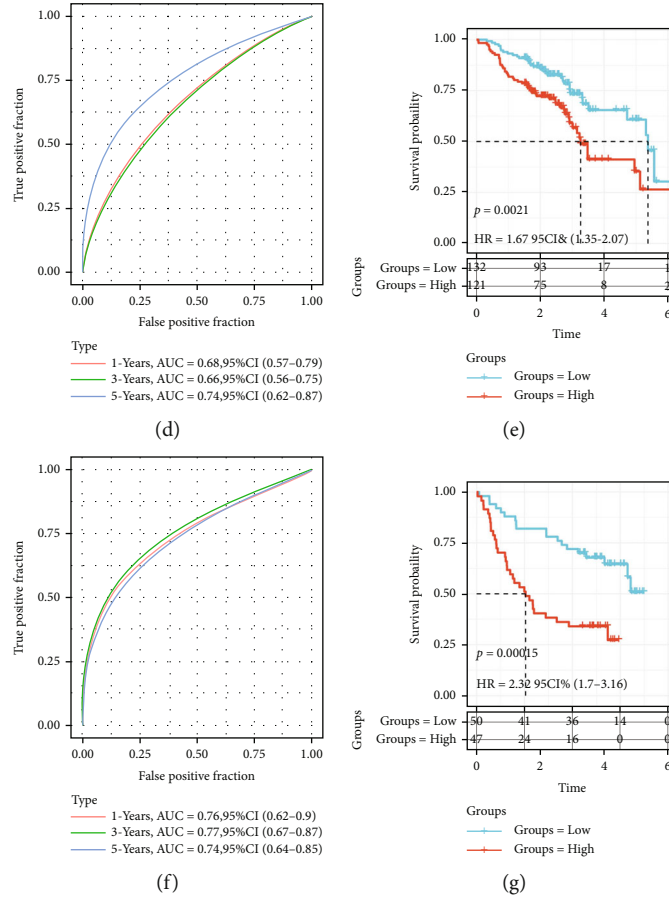
FIGURE 5: Continued.

(d)

(e)

(f)

(g)

FIGURE 5: Clinical prognostic modeling and validation. (a) The distribution of survival status, CSRS.Score, and survival time corresponding to senescence-related genes expression in the TCGA cohort. (b) ROC curve with AUC for CSRS.Score classification in the TCGA cohort. (c) KM survival curves of two risk groups in the TCGA cohort. (d, e) ROC curves and KM survival curves of CSRS.Score in the GSE65858 cohort. (f, g) ROC curves and KM survival curves of CSRS.Score in the GSE41613 cohort.

performance of CSRS.Score in different clinical features, samples were also clearly divided into two risk groups with differential overall survival ($P < 0.0001$, Figure S3).

3.6. CSRS.Score Was Correlated with Immune Infiltration and Immunotherapy/Chemotherapy Responses.

To further explore the immune characteristics between two risk groups, we analyzed the estimated proportion of 22 immune cells in high- and low-risk groups in the TCGA cohort (Figure 6(a)). Some immune cells were differently enriched in high- and low-CSRS.Score subgroups, such as M1 macrophages, activated CD4 memory T cells, and CD8 T cells. Overall, the high-risk group had lower immune infiltration than the low-risk group according to ESTI-MATE analysis (Figure 6(b)). The correlation analysis between CSRS.Score and 22 immune cells demonstrated that CSRS.Score was significantly correlated with resting CD4 memory T cells, M0 macrophages, mast cells, CD8 T cells, follicular helper T cells, and regulatory T cells ($P < 0.001$, Figure 6(c)).

We further explored the immune response of two risk groups to immunotherapy. The expression of most immune checkpoints was differential between high- and low-risk groups (Figure 6(d)). TIDE prediction showed that the high-risk group had higher scores of two immunosuppressive cells (MDSC and CAF) probably contributing to higher T cell exclusion score and TIDE score (Figure 6(e) and Figure S4A-B), indicating that the high-risk group was more liable to escape from immunotherapy. In addition to immunotherapeutic response, we also assessed the response of two risk groups to chemotherapeutic drugs (paclitaxel, docetaxel, cisplatin, and 5-fluorouracil). The results showed that the high-risk group was more sensitive to paclitaxel, docetaxel, cisplatin, and 5-fluorouracil (Figure S4C-E).

3.7. CSRS.Score Incorporates Clinicopathological Features to Further Improve Survival Prediction.

We constructed a decision tree based on clinical information and CSRS.Score, and only stage, gender, and CSRS.Score were remained in the decision tree (Figure 7(a)). Four subgroups including low, median, high, and highest were determined with differential overall survival (Figure 7(b)). Only highest subgroup consisted of high-risk samples (Figure 7(c)) and highest subgroup had the most proportions of C1 and C2 subtypes (Figure 7(d)). Univariate and multifactorial Cox regression analyses illustrated that CSRS.Score was the most significant prognostic factor and age and stage were also independent risk factors (Figures 7(e) and 7(f)). Then, we established a
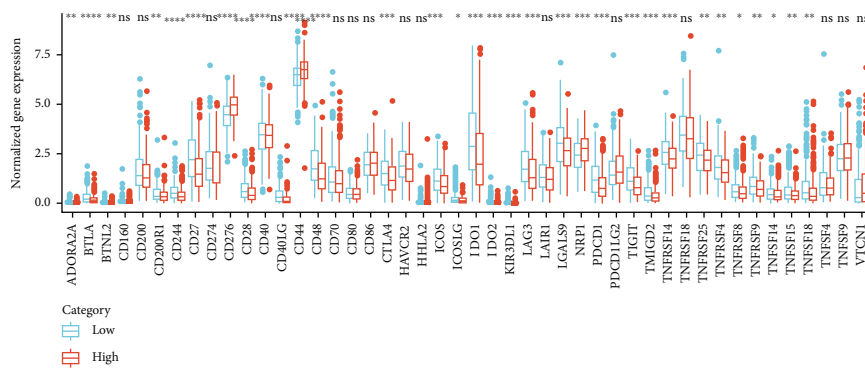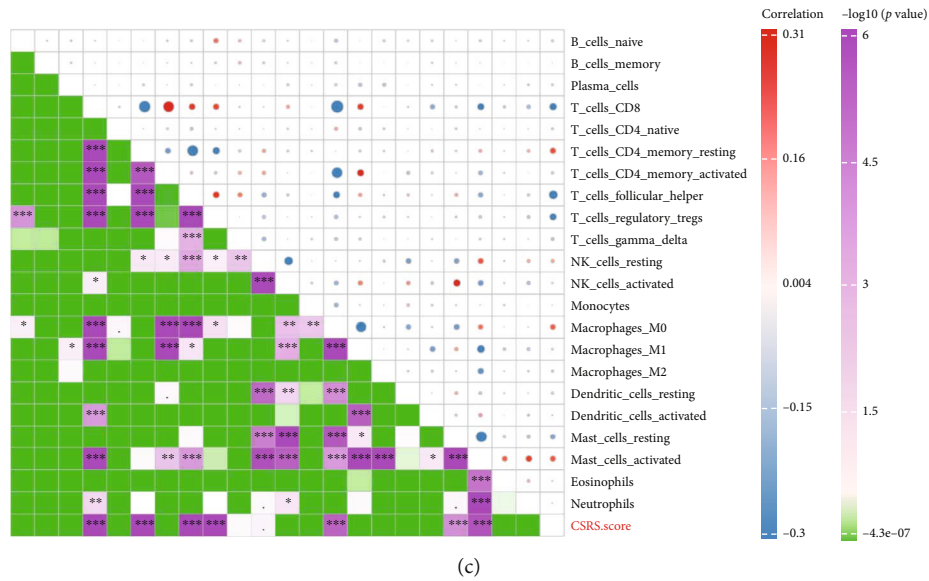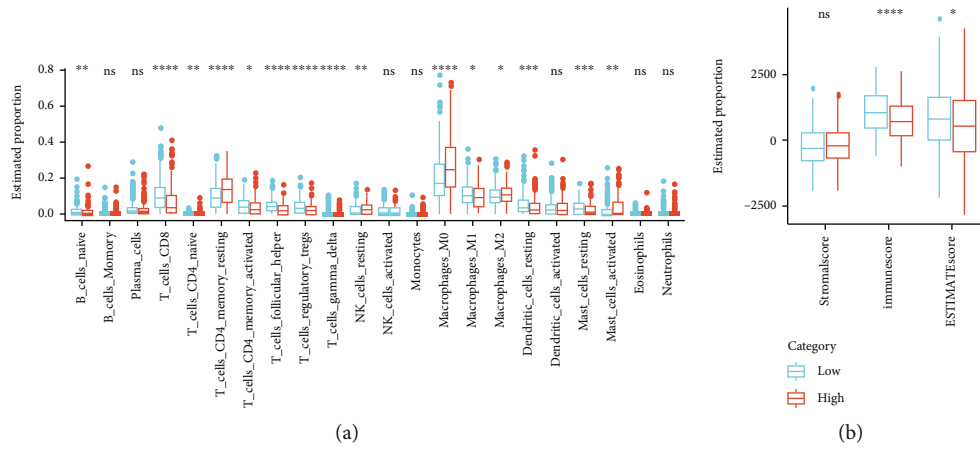
(a)

(b)
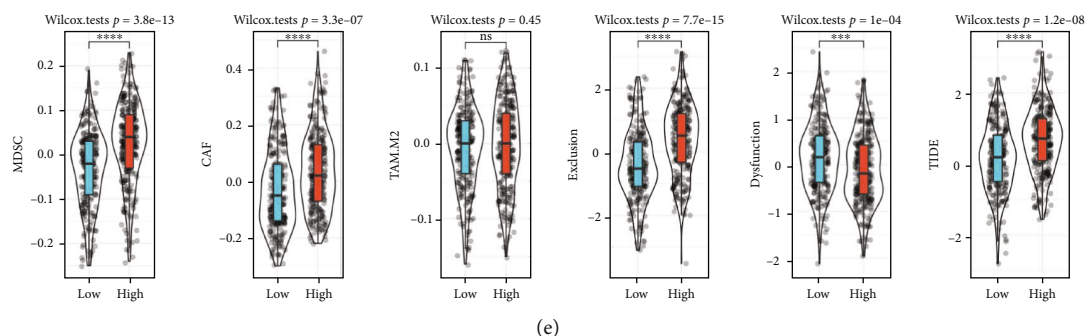

(c)

(d)

FIGURE 6: Continued.

(e)

FIGURE 6: Immune characteristics between two risk groups in the TCGA cohort. (a) The quantity of 22 immune cells analyzed by CIBERSORT. (b) The TCGA cohort immune score and stromal score evaluated by ESTIMATE software. (c) Correlation analysis on CSRS. Score and 22 immune cells in the TCGA cohort. (d) The level of immune checkpoint genes in two risk groups. (e) TIDE analysis for predicting immune response to immune checkpoint blockade. ns: no significance. $^{****}P < 0.0001$, $^{***}P < 0.001$, $^{**}P < 0.01$, and $^{*}P < 0.05$.

nomogram based on the three independent risk factors for predicting 1-year, 3-year, and 5-year survival (Figure 7(g)). The predictive accuracy of the nomogram was validated by calibration curve (Figure 7(h)). The predicted overall survival of 1 year, 3 years, and 5 years fitted with the observed overall survival, indicating the reliability of the nomogram. Decision curve analysis showed that the nomogram reached relatively high net benefit compared with others (Figure 7(i)). Both nomogram and CSRS.Score had a higher AUC than other clinical features (Figure 7(j)), proving a favorable performance for predicting prognosis in HNSCC patients.

4. Discussion

A number of evidences prove that senescence-dependent changes are associated with proinflammatory properties and are involved in the chronic inflammatory microenvironment. Increased levels of inflammation and immune cell infiltration with senescence-dependent changes may lead to tumor formation and malignancy, and their function in HNC is unknown [34]. Exploring the molecular mechanisms of cellular senescence-associated genes is important to determine the role of senescence-dependent changes in HNC. To date, few studies have systematically investigated the molecular mechanisms of cellular senescence-associated genes in HNC and the association between cellular senescence-associated genes and HNC prognosis. We used cellular senescence-associated genes to identify four molecular subtypes by consensus clustering, and these four molecular subtypes differed significantly in prognosis and several clinical features.

Senescence and tumor microenvironment are closely related in tumor progression and invasion. There is an obvious difference of immunity between elder patients and younger patients, where younger patients have more abundant T cells in tumor tissue than elder patients. Senescence leads to a declining immune system which is referred to as immune senescence [35]. It is suggested that tumor-infiltrating CD4+ and CD8+ T cells are reduced in old mice compared to young mice and old mice are more infiltrated with neutrophils and macrophages [36]. Our results showed that CD8 T cells and activated memory CD4 T cells were lower enriched in C1 subtype, while M0 macrophages were extremely higher enriched. C1 also performed a lower stromal and immune infiltration than other subtypes. GSEA results revealed that the immune-related pathways were suppressed in the C1 subtype, which was consistent with its immune features. Therefore, we inferred that cellular senescence-associated genes may play a large role in immune-related pathways and tumor infiltration.

Based on the DEGs among different subtypes, we confirmed a total of eight key cellular senescence prognosis-related genes including CDKN2A, PYGL, KRT8, AREG, MAGEA4, DES, EPHX3, and SPINK6. CDKN2A has been shown to mediate the antitumor effects in HNSCC through cell cycle arrest [37]. Low CDKN2A expression predicts unfavorable prognosis in HPV-negative HNSCC independent of other clinical factors [38], which is accordant with our result that CDKN2A is lower expressed in high-risk group. PYGL is significantly associated with overall survival in HNC patients and may be an independent risk factor for HNSCC prognosis [39]. Keratin 8 (KRT8) overexpression enhanced cell proliferation and migration in gastric cancer and lung cancer, while its decreased expression markedly inhibited cell proliferation, migration, and EMT process [40, 41]. A multiscale integrated analysis figured out KRT8 as a pan-cancer early biomarker [42]. Amphiregulin (AREG) is a ligand of epidermal growth factor receptor (EGFR), which is underlined to function in several aspects of cancerogenesis including cancer cell growth, invasion, metastasis, angiogenesis, and resistance to apoptosis [43].

Additionally, the correlation of molecular subtypes with gene mutations was analyzed, and a significant correlation between the two was detected. The common TP53 and CDKN2A were mutated at a high frequency in the four subtypes. We further evaluated the degree of clinical response of conventional chemotherapeutic drugs paclitaxel, docetaxel, cisplatin, and 5-fluorouracil to CSRS.Score subgroups, and the results showed that high-CSRS.Score to paclitaxel, docetaxel, cisplatin, and 5-fluorouracil was more sensitive. The results suggest that different subgroups have different degrees corresponding to different chemotherapeutic drugs, and perhaps our screened aging-related genes can be used as biomarkers of clinical drug treatment response.
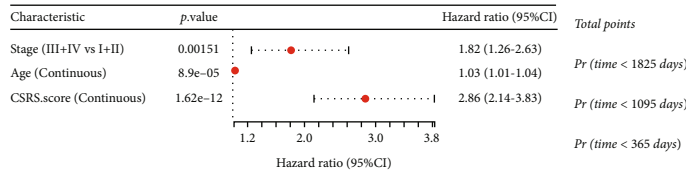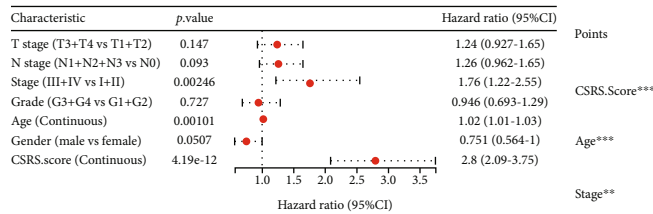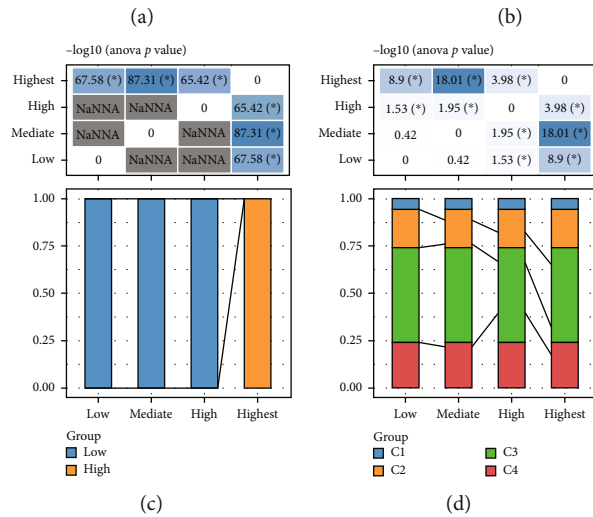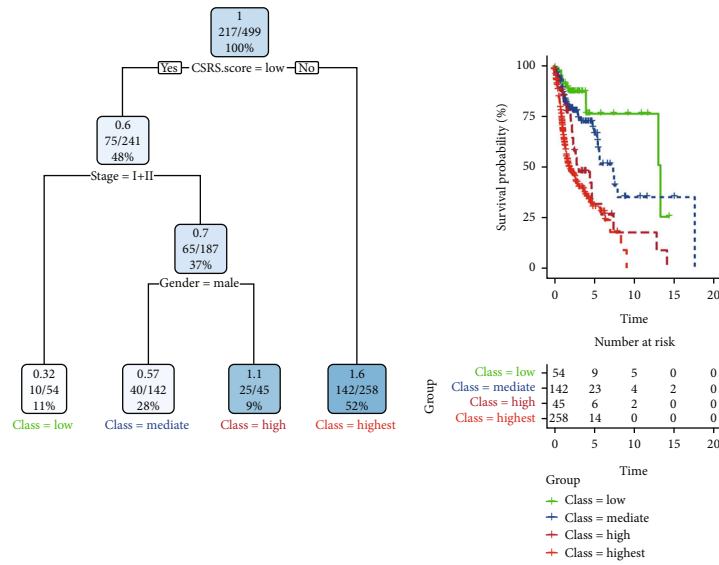
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 7: Continued.
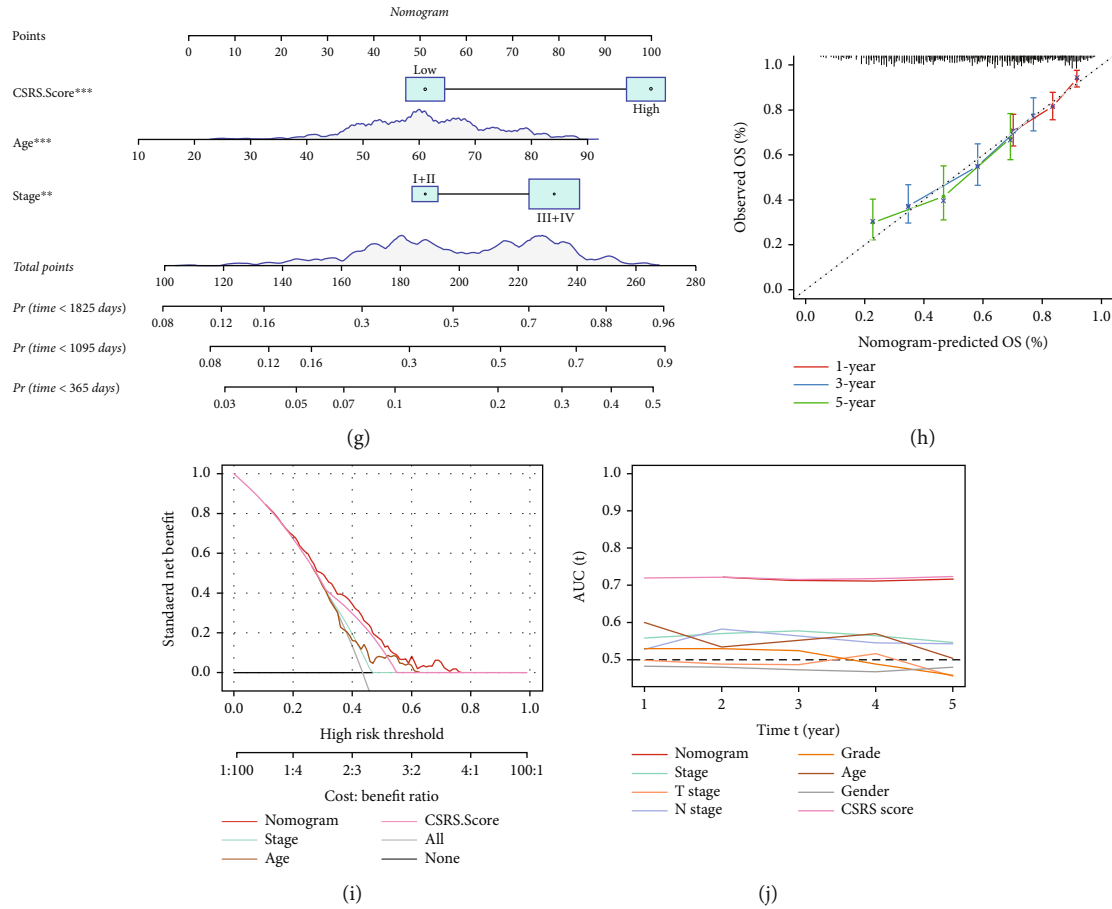
(g)

(h)

(i)

(j)

FIGURE 7: CSRS.Score combined with clinicopathological characteristics to further improve prognostic models and survival prediction. (a) Full-scale annotations of patients including gender, CSRS.Score, age, grade, and TNM stage were applied to develop a survival decision tree for optimizing the risk stratification. (b) Among the four subgroups, significant differences in overall survival could be found. (c, d) Comparative analysis between different subgroups. (e, f) Univariate (e) and multivariate (f) Cox analysis of CSRS.Score and clinicopathological features. (g) A nomogram was developed based on age, CSRS.Score, and stage. (h) Calibration curves for 1, 3, and 5 years for the columnar graph. (i) Decision curve analysis of the nomogram, CSRS.Score, and other clinical features. (j) Compared with other clinicopathological features, the nomogram exhibited the most powerful capacity for survival prediction.

Although the prognostic risk model based senescence-associated genes was illustrated to have robust performance in several independent cohorts, this study did not validate the model in the wet experiments. More clinical HNSCC samples should be included to support the reliability and accuracy of the eight prognostic biomarkers in the future.

5. Conclusion

In this study, we developed a prognostic risk model with cellular senescence-associated genes that has great potential as a biomarker for HNSCC patients and provides insights into individualized immunotherapy for head and neck cancer patients.

Data Availability

The datasets analyzed in this study could be found in GSE65858 at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65858 and GSE41613 at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41613.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Rensheng Wang and Lin Deng conceived and designed the research. Jinglin Mi and Guozhen Zhang acquired the data. Xiaolan Ruan and Lin Deng analyzed and interpreted the data. Yufei Pan carried out statistical analysis. Yufei Pan drafted the manuscript. Rensheng Wang revised the manuscript for important intellectual content. Lin Deng and Jinglin Mi are equal contributors and co-first authors.

Supplementary Materials

Figure S1: clinicopathological characteristics between molecular subtypes. (A–I) Clinicopathological characteristics of molecular subtypes in the TCGA-HNSC cohort, where the lower half is the proportion and the upper half is the statistical significance of the difference in distribution between the

two-log10 (P value). Figure S2: results of GSEA analysis of pathways between different molecular subtypes in the TCGA cohort. Figure S3: KM curves between high- and low-risk groups of CSRS.Score in different clinical features in the TCGA cohort. Figure S4: response differences to immuno-therapy/chemotherapy in two risk groups. (A, B) TIDE analysis results between two risk groups in the GSE65858 (A) and GSE41613 (B) cohorts. (C–E) The box plots of the estimated IC50 for paclitaxel, docetaxel, cisplatin, and 5-fluorouracil in the TCGA cohort, GSE65858, and GSE41613 cohorts. Table S1: a list of 28 prognostic cellular senescence-associated genes used for molecular subtyping. (Supplementary Materials)

References

[1] D. M. Shin, S. Nannapaneni, M. R. Patel et al., "Phase Ib study of chemoprevention with green tea polyphenon E and erlotinib in patients with advanced premalignant lesions (APL) of the head and neck," *Clinical Cancer Research*, vol. 26, no. 22, pp. 5860–5868, 2020.

[2] S. Marur and A. A. Forastiere, "Head and neck cancer: changing epidemiology, diagnosis, and treatment," *Mayo Clinic Proceedings*, vol. 83, no. 4, pp. 489–501, 2008.

[3] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[4] N. Nishio, N. S. van den Berg, S. van Keulen et al., "Optical molecular imaging can differentiate metastatic from benign lymph nodes in head and neck cancer," *Nature Communications*, vol. 10, no. 1, p. 5044, 2019.

[5] T. Nakashima, R. Yasumatsu, K. Asai, H. Uryu, R. Kogo, and T. Nakagawa, "Single-cycle induction chemotherapy for resectable advanced hypopharyngeal cancer," *International Journal of Clinical Oncology*, vol. 22, no. 3, pp. 442–447, 2017.

[6] F. J. Civantos, R. P. Zitsch, D. E. Schuller et al., "Sentinel lymph node biopsy accurately stages the regional lymph nodes for T1-T2 oral squamous cell carcinomas: results of a prospective multi-institutional trial," *Journal of Clinical Oncology*, vol. 28, no. 8, pp. 1395–1400, 2010.

[7] J. Campisi, "Aging, cellular senescence, and cancer," *Annual Review of Physiology*, vol. 75, no. 1, pp. 685–705, 2013.

[8] P. A. Pérez-Mancera, A. R. Young, and M. Narita, "Inside and out: the activities of senescence in cancer," *Nature Reviews Cancer*, vol. 14, no. 8, pp. 547–558, 2014.

[9] M. K. Ruhland, A. J. Loza, A. H. Capietto et al., "Stromal senescence establishes an immunosuppressive microenvironment that drives tumorigenesis," *Nature Communications*, vol. 7, no. 1, article 11762, 2016.

[10] L. Yuan, L. Zhai, L. Qian et al., "Switching off IMMP2L signaling drives senescence via simultaneous metabolic alteration and blockage of cell death," *Cell Research*, vol. 28, no. 6, pp. 625–643, 2018.

[11] M. Xu, T. Pirtskhalava, J. N. Farr et al., "Senolytics improve physical function and increase lifespan in old age," *Nature Medicine*, vol. 24, no. 8, pp. 1246–1256, 2018.

[12] N. Tasdemir, A. Banito, J. S. Roe et al., "BRD4 connects enhancer remodeling to senescence immune surveillance," *Cancer Discovery*, vol. 6, no. 6, pp. 612–629, 2016.

[13] W. Xie, I. Kagiampakis, L. Pan et al., "DNA methylation patterns separate senescence from transformation potential and indicate cancer risk," *Cancer Cell*, vol. 33, no. 2, pp. 309–321.e5, 2018.

[14] M. M. Sugrue, D. Y. Shin, S. W. Lee, and S. A. Aaronson, "Wild-type p 53 triggers a rapid senescence program in human tumor cells lacking functional p 53," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 18, pp. 9648–9653, 1997.

[15] R. Haq, J. D. Brenton, M. Takahashi et al., "Constitutive p38HOG mitogen-activated protein kinase activation induces permanent cell cycle arrest and senescence," *Cancer Research*, vol. 62, no. 17, pp. 5076–5082, 2002.

[16] A. Alimonti, C. Nardella, Z. Chen et al., "A novel type of cellular senescence that can be enhanced in mouse models and human tumor xenografts to suppress prostate tumorigenesis," *The Journal of Clinical Investigation*, vol. 120, no. 3, pp. 681–693, 2010.

[17] D. A. Gewirtz, S. E. Holt, and L. W. Elmore, "Accelerated senescence: an emerging role in tumor cell response to chemotherapy and radiation," *Biochemical Pharmacology*, vol. 76, no. 8, pp. 947–957, 2008.

[18] P. Hari, F. R. Millar, N. Tarrats et al., "The innate immune sensor Toll-like receptor 2 controls the senescence-associated secretory phenotype," *Science Advances*, vol. 5, no. 6, p. eaaw0254, 2019.

[19] B. Zhao, P. Liu, T. Fukumoto et al., "Topoisomerase 1 cleavage complex enables pattern recognition and inflammation during senescence," *Nature Communications*, vol. 11, no. 1, p. 908, 2020.

[20] J. A. Ewald, J. A. Desotelle, G. Wilding, and D. F. Jarrard, "Therapy-induced senescence in cancer," *Journal of the National Cancer Institute*, vol. 102, no. 20, pp. 1536–1546, 2010.

[21] M. Althubiti, L. Lezina, S. Carrera et al., "Characterization of novel markers of senescence and their prognostic potential in cancer," *Cell Death & Disease*, vol. 5, no. 11, p. e1528, 2014.

[22] D. Coppola, L. Balducci, D. T. Chen et al., "Senescence-associated-gene signature identifies genes linked to age, prognosis, and progression of human gliomas," *Journal of Geriatric Oncology*, vol. 5, no. 4, pp. 389–399, 2014.

[23] J. Yang, Q. Jiang, L. Liu et al., "Identification of prognostic aging-related genes associated with immunosuppression and inflammation in head and neck squamous cell carcinoma," *Aging*, vol. 12, no. 24, pp. 25778–25804, 2020.

[24] M. D. Wilkerson and D. N. Hayes, "ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, pp. 1572-1573, 2010.

[25] M. E. Ritchie, B. Phipson, D. Wu et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, article e47, 2015.

[26] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.

[27] Z. Zhang, "Variable selection with stepwise and best subset approaches," *Annals of Translational Medicine*, vol. 4, no. 7, p. 136, 2016.

[28] B. Chen, M. S. Khodadoust, C. L. Liu, A. M. Newman, and A. A. Alizadeh, "Profiling tumor infiltrating immune cells with

CIBERSORT," *Methods in Molecular Biology*, vol. 1711, pp. 243–259, 2018.

[29] K. Yoshihara, M. Shahmoradgoli, E. Martínez et al., "Inferring tumour purity and stromal and immune cell admixture from expression data," *Nature Communications*, vol. 4, no. 1, p. 2612, 2013.

[30] P. Jiang, S. Gu, D. Pan et al., "Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response," *Nature Medicine*, vol. 24, no. 10, pp. 1550–1558, 2018.

[31] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, pp. 15545–15550, 2005.

[32] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, and P. Tamayo, "The Molecular Signatures Database (MSigDB) hallmark gene set collection," *Cell Systems*, vol. 1, no. 6, pp. 417–425, 2015.

[33] V. Thorsson, D. L. Gibbs, S. D. Brown et al., "The immune landscape of cancer," *Immunity*, vol. 48, no. 4, pp. 812–30.e14, 2018.

[34] T. Kuilman, C. Michaloglou, L. C. Vredeveld et al., "Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network," *Cell*, vol. 133, no. 6, pp. 1019–1031, 2008.

[35] R. M. Gorczynski and E. Terzioglu, "Aging and the immune system," *International Urology and Nephrology*, vol. 40, no. 4, pp. 1117–1125, 2008.

[36] M. Provinciali, K. Argentati, and A. Tibaldi, "Efficacy of cancer gene therapy in aging: adenocarcinoma cells engineered to release IL-2 are rejected but do not induce tumor specific immune memory in old mice," *Gene Therapy*, vol. 7, no. 7, pp. 624–632, 2000.

[37] S. Bhattacharyya, V. Sekar, B. Majumder et al., "CDKN2A-p53 mediated antitumor effect of lupeol in head and neck cancer," *Cellular Oncology*, vol. 40, no. 2, pp. 145–155, 2017.

[38] W. S. Chen, R. S. Bindra, A. Mo et al., "CDKN2A copy number loss is an independent prognostic factor in HPV-negative head and neck squamous cell carcinoma," *Frontiers in Oncology*, vol. 8, p. 95, 2018.

[39] L. Chen, X. He, S. Yi, G. Liu, Y. Liu, and Y. Ling, "Six glycolysis-related genes as prognostic risk markers can predict the prognosis of patients with head and neck squamous cell carcinoma," *BioMed Research International*, vol. 2021, Article ID 8824195, 13 pages, 2021.

[40] H. Chen, X. Chen, B. Pan, C. Zheng, L. Hong, and W. Han, "KRT8 serves as a novel biomarker for LUAD and promotes metastasis and EMT via NF-κB signaling," *Frontiers in Oncology*, vol. 12, article 875146, 2022.

[41] J. Fang, H. Wang, Y. Liu, F. Ding, Y. Ni, and S. Shao, "High KRT8 expression promotes tumor progression and metastasis of gastric cancer," *Cancer Science*, vol. 108, no. 2, pp. 178–186, 2017.

[42] M. K. D. Scott, M. Limaye, S. Schaffert et al., "A multi-scale integrated analysis identifies KRT8 as a pan-cancer early biomarker," *Pacific Symposium on Biocomputing*, vol. 26, pp. 297–308, 2021.

[43] B. Busser, L. Sancey, E. Brambilla, J. L. Coll, and A. Hurbin, "The multiple roles of amphiregulin in human cancer," *Biochimica et Biophysica Acta*, vol. 1816, no. 2, pp. 119–131, 2011.