Check for updates

OPEN

Characterization of macrophages in head and neck squamous cell carcinoma and development of MRG-based risk signature

Lei Liu & Qiang Liu✉

Macrophages are immune cells in the TME that can not only inhibit angiogenesis, extracellular matrix remodeling, cancer cell proliferation, and metastasis but also mediate the phagocytosis and killing of cancer cells after activation, making them key targets in anti-tumor immunotherapy. However, there is little research on macrophages and their relation to disease prognosis in HNSCC. Initially, we collected scRNA-seq, bulk RNA-seq, and clinical data. Subsequently, we identified macrophages and distinguished MRGs. Using the K-means algorithm, we performed consensus unsupervised clustering. Next, we used ssGSEA analysis to assess immune cell infiltration in MRG clusters. A risk model was established using multivariate Cox analysis. Then, Kaplan–Meier, ROC curves, univariate and multivariate COX analyses, and C-index was used to validate the predictive power of the signature. The TIDE method was applied to assess the response to immunotherapy in patients diagnosed with HNSCC. In addition, drug susceptibility predictions were made for the GDSC database using the calcPhenotype function. We found that 8 MRGs had prognostic potential. Patients in the MRG group A had a higher probability of survival, and MRG clusters A and B had different characteristics. Cluster A had a higher degree of expression and infiltration in MRG, indicating a closer relationship with MRG. The accuracy of the signature was validated using univariate and multivariate Cox analysis, C-index, and nomogram. Immune landscape analysis found that various immune functions were highly expressed in the low-risk group, indicating an improved response to immunotherapy. Finally, drugs with high sensitivity to HNSCC (such as 5-Fluorouracil, Temozolomide, Carmustine, and EPZ5676) were explored and analyze the malignant characteristics of HNSCC. We constructed a prognostic model using multivariate Cox analysis, consisting of 8 MRGs (TGM2, STC1, SH2D3C, PIK3R3, MAP3K8, ITGA5, ARHGAP4, and AQP1). Patients in the low-risk group may have a higher response to immunotherapy. The more prominent drugs for drug selection are 5-fluorouracil, temozolomide and so on. Malignant features associated with HNSCC include angiogenesis, EMT, and the cell cycle. This study has opened up new prospects for the prognosis, prediction, and clinical treatment strategy of HNSCC.

Head and neck squamous cell carcinoma (HNSCC) is the main type (more than 90%) of head and neck tumors (common sites: mouth, pharynx, tongue, larynx, etc.)[1–3]. The major risk factors for the disease are smoking, alcohol abuse, local repeated irritation, viral infection, and so on[3–6]. Traditional surgery, radiotherapy, and chemotherapy are still important comprehensive treatment options for HNSCC, but with the fast progress of targeted therapy and immunotherapy study, the comprehensive treatment of this disease (especially advanced HNSCC) has gradually approached immunotherapy[7–9]. In previous studies, targeted immunotherapy increased survival in patients with HNSCC, but fewer than one in five patients had a long-lasting treatment response[10–13]. Given this situation, we believe that investigating new immunotherapies by studying key players in the tumor microenvironment (TME) of HNSCC can further improve the immune response of tumor cells and, accordingly, the therapeutic efficacy of HNSCC.

Department of Otorhinolaryngology, The Affiliated Traditional Chinese Medicine Hospital, Southwest Medical University, Luzhou 646000, China. ✉email: lq@swmu.edu.cn

nature portfolio

1

Macrophages are derived from monocytes, which are a type of white blood cell. As the name implies, tumor-associated macrophages (TAMs) are immune cells present in the TME. They can inhibit angiogenesis, extracellular matrix remodeling, cancer cell proliferation, and metastasis. Additionally, they can mediate the phagocytosis and killing of cancer cells after activation, so they are key targets in anti-tumor immunotherapy[14]. Macrophage-related genes (MRGs) can determine the morphology and function of macrophages, and MRGs can exert an essential function in tumor recognition, phagocytosis, and cytokine release by regulating macrophages, which are important in controlling and eliminating tumors[15,16]. Exosomes can activate the NF-κB pathway in macrophages, thereby enhancing the expression of pro-inflammatory elements, thereby promoting the proliferation and metastasis of gastric cancer cells[17]. The risk scoring system formed by using 11 MRGs predicts the prognosis of lung cancer patients and evaluates their immune response. State of infiltration[18]. In colon cancer, four central genes were recognized as associated with M2 macrophages, and the M2I score was established accordingly, among which TMB, MSI, and sensitivity are higher[19]. There are countless similar studies that have demonstrated that macrophages, which are regulated by macrophage-related genes, can play an important in tumor regulation and prognosis.

The literature has pointed out that in HNSCC, TAM is related to poor prognosis because it promotes tumor progression and suppresses immune response through innate and adaptive immune mechanisms, but TAM as a double-edged sword can also inhibit tumor cells by depleting immunosuppressive function or stimulating anti-tumor ability[20]. In this study, the cell data of HNSCC will be obtained through public datasets, and then the data will be analyzed to obtain "MRGs that can judge prognosis" and "disease prognosis prediction models based on MRGs", so as to provide more scientific medical suggestions for the intervention of HNSCC patients.

Methods

Preparation of data

This study utilized multiple publicly available datasets to investigate the molecular mechanisms underlying HNSCC. Specifically, scRNA-seq data from 18 HNSCC samples were obtained from the GEO-GSE103322 dataset, resulting in a total of 5902 single cells being analyzed[21]. RNA-seq and clinical data for HNSCC were also obtained from the TCGA-HNSCC, GSE27020, GSE41613, and GSE65858 datasets. After removing incomplete and duplicate data, the study analyzed 511 HNSCC samples from TCGA-HNSCC, 109 from GEO-GSE27020, 97 from GEO-GSE41613, and 270 from GEO-GSE65858[22–25]. Additionally, RNA-seq and clinical data were obtained from 8739 samples representing 32 different types of tumors from USCS Xena[22–25]. Figure 1 illustrates the general workflow of this research.

Manipulation of scRNA-seq data

Firstly, we implemented the "Seurat" package to create a Seurat object, which served as the foundation for analyzing the scRNA-seq data[27]. We conducted quality control procedures based on thresholds established in previous studies[21]. Subsequently, we processed cell cycle effects, performed data normalization, employed dimensionality reduction techniques (1:30), proceeded with clustering analysis (resolution = 0.5), and allocated
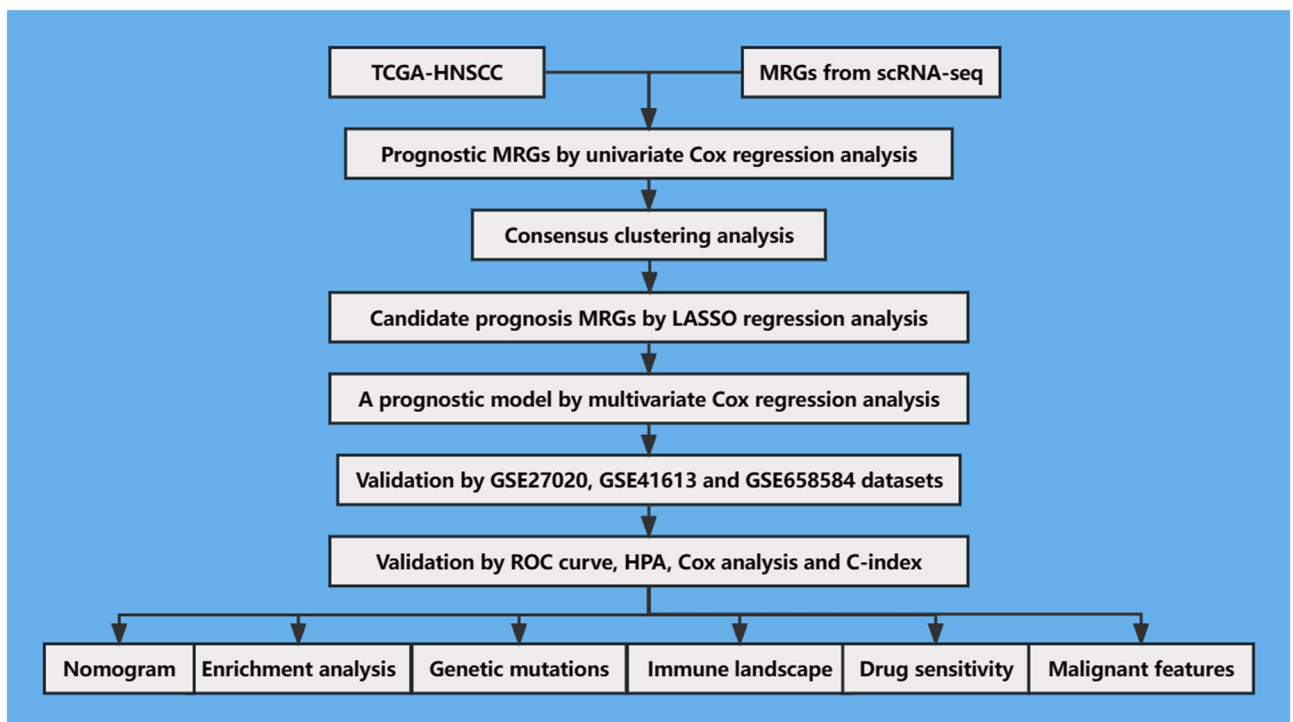
Figure 1. The overall workflow of the study.

cell annotations[28,29]. To recognize MRGs, we employed the "FindAllMarkers" function, which allowed us to recognize highly variable genes meeting specific criteria (logFC ≥ 0.3, min.pct = 0.3, and diff.pct ≥ 0.2) within the macrophage subset. These genes were deemed as MRGs, denoting their potential involvement in macrophage function and biology. The utilization of the Seurat package, in combination with established guidelines and statistical approaches, ensures the robustness and validity of the results.

Recognition of prognostic MRGs

To investigate the prognostic significance of MRGs in HNSCC, we performed univariate Cox analysis ($p < 0.05$), and aimed at identifying MRGs significantly associated with survival outcomes in HNSCC. After this step, we employed a comprehensive analytical strategy that involved an in-depth examination of the inter-expression relationship between the CNV landscape and the recognized prognostic MRGs[30,31].

Consensus clustering analysis

To elucidate the clinically and pathologically relevant aspects of the recognized MRGs, we utilized the K-means algorithm with consensus unsupervised clustering analysis through the R package "ConsensusClusterPlus"[32]. This method allowed us to identify different clusters of samples according to the expression levels of the MRGs, with the selection of optimal K values[33]. We evaluated differences in survival outcomes among the MRG clusters with survival analyses, enabling us to determine the potential clinical significance of these MRG clusters. To explore the distribution of MRG clusters, we employed dimensionality reduction techniques, including PCA, tSNE, and UMAP analyses. Besides, we evaluated the relationship between MRG clusters and specific clinical features. In summary, these analytical approaches enabled us to gain a comprehensive understanding of the potential significance of the recognized MRGs, with the potential to inform tailored therapeutic interventions and improve clinical outcomes for HNSCC patients.

Exploration of immune status and enrichment analysis

Firstly, we conducted an analysis of the expression patterns of MRGs among MRG clusters, aiming to recognize distinct regulatory patterns among MRGs. By examining the expression patterns, we intended to obtain a comprehensive understanding of the different regulatory mechanisms at play in HNSCC. Additionally, we used ssGSEA to assess the levels of immune cell infiltration in MRG clusters[34]. We sought to discern any significant differences in immune cell infiltration between the clusters with the Wilcoxon test. This method would provide valuable information regarding the potential influence of prognostic MRGs on the TME and immune response in HNSCC. Furthermore, we implemented the GSVA method to assess the activation status of KEGG pathways[34]. By employing this algorithm, we aimed to identify the key pathways related to each cluster. To visualize the results and highlight the most significant enrichment outcomes, we generated a heat map displaying the top 30 enrichment results.

Construction of prognostic MRG model

We implemented LASSO regression analysis to screen the candidate MRGs and diminish any irrelevant and redundant factors. By introducing a penalty term to the conventional regression model, the LASSO analysis allowed us to reduce the complexity of the signature by shrinking the coefficients of unimportant genes to zero while keeping the most significant genes[35]. Furthermore, we applied multivariate Cox analysis to develop a prognostic model[36].

Validation of prognostic MRG model

Survival analysis was applied to examine the differences in survival. Furthermore, the accuracy of the prognostic signature was assessed using ROC curves, which allowed for an evaluation of its predictive performance over time[37]. To ensure the generalizability of the model, HNSCC samples from multiple datasets, including GSE27020, GSE41613, and GSE65858, were utilized for external validation. The validity of the model was further examined using the HPA database, which provided information on the protein expression levels of the identified prognostic MRGs in both normal and HNSCC tissues. This analysis offered worthwhile insights into the potential utility of the model in guiding treatment decisions for specific patient populations. Moreover, both univariate and multivariate Cox analyses were conducted to assess the independent predictive power of the prognostic model. The C-index was employed to quantify the discriminatory power of the signature with respect to traditional clinical features. This assessment provided a quantitative measure of the model's predictive performance and contributed to the determination of its added value beyond established prognostic factors. Lastly, a nomogram, incorporating both the clinical data and the signature, was developed to predict the survival rates of HNSCC. By incorporating multiple prognostic factors into a single model, the nomogram provided a practical tool for individualized patient prognosis estimation.

Enrichment and gene mutation analyses

This study employed advanced methodologies to identify DEGs across distinct risk groups. To ensure the robustness of the findings, DEGs were identified using stringent criteria, including a threshold of |logFC > = 1.5| and a FDR of less than 0.05. Subsequently, the identified DEGs underwent GO and KEGG analyses ($p < 0.05$), to elucidate their biological functions and involve signaling pathways[38–41]. To further explore the genomic landscape of HNSCC and its association with different risk scores, the "Maftools" package was used[42]. This tool enabled the identification and comparison of the mutation burden between various risk score groups. By quantifying the

number of mutations present, the implication of genomic alterations in contributing to different levels of risk in HNSCC could be investigated.

Exploration of immunization status

We employed a range of analytical approaches to assess immune cell infiltration. Various methods were used to deconvolve and quantify the immune cell composition in the TME[43–49]. The ssGSEA approach was employed to investigate the immune function of the risk groups. This method allowed for the quantification of immune function scores, which reflect the overall immune activity and response within each risk group. Moreover, the expression levels of ICGs within the identified risk groups were examined. ICGs play a crucial role in regulating the immune response and can impact the success of immunotherapeutic interventions. The TIDE method was employed to forecast the immunotherapy response[50]. This algorithm takes into account various immune-related factors, including tumor immune evasion and cytotoxic immune infiltration, to forecast the likelihood of an immunotherapeutic response.

Recognition of medicines

To investigate the efficacy of different drugs against HNSCC, we utilized the "calcPhenotype" function in the "oncoPredict" package to forecast drug sensitivity[51]. By applying the approach to the GDSC database, the study was able to recognize the efficacy of specific drugs for HNSCC.

Correlation with malignant features

The z-score method is used to normalize gene expression data, facilitates comparative analyses between different samples, and enables calculation of the activity of specific pathways and integration of the expression of characteristic genes[26,52]. Based on the "GSVA" package, we explored the relationship between the MRG signature and angiogenesis, cell cycle regulation, and EMT by comparing them with the pearson algorithm[53,54].

Result

Recognition of MRGs

This study investigated the molecular characteristics of HNSCC using scRNA-seq analysis. 18 HNSCC samples were collected, and a dataset of 5902 cells from the GSE103322 database was analyzed after quality control (Fig. 2A). The scRNA-seq data were normalized, reduced dimensionality, and clustered using tSNE plot analysis, revealing the distribution of different cell populations (Fig. 2B). Cell annotations for each cluster were performed (Fig. 2C) and bubble plots were generated to demonstrate the expression of genes (Fig. 2D). Additionally, the "FindAllMarkers" function was utilized to recognize 216 significantly differentially expressed MRGs (Table S1).

Recognition of prognostic MRGs

We conducted a univariate Cox analysis, showing that 33 MRGs were significantly related to overall survival (Fig. 3A). To understand the potential genetic alterations affecting these MRGs, the locations of CNVs on chromosomes were tagged (Fig. 3B). A network analysis was performed to explore the interconnections and regulatory relationships among the MRGs, providing a comprehensive understanding of their biological functions and potential interactions, as depicted in Fig. 3C. Furthermore, the CNVs of the 33 identified MRGs were investigated. Among these genes, ARRB1, ARHGAP4, CLIC2, SOCS3, ICAM2, VAMP5, PLVAP, PIK3R3, SH2D3C, LGALS9, CIITA, FXYD5, CD34, AQP1, SELP, MAP3K8, RASGRP3, and RAMP3 showed a higher frequency of amplification CNVs, indicating potential oncogenic roles for these genes. On the other hand, LI1, RASSF2, IL3RA, FGD2, IRF8, GIMAP4, GIMAP6, ITGA5, TGM2, RAMP2, CSF2RB, ARHGAP29, PALMD, MMRN2, and STC1 exhibited a higher frequency of missing CNVs for deletions, suggesting potential tumor suppressor roles for these genes (Fig. 3D).

Recognition of MRG clusters

This study employed the consensus clustering approach to divide patients into distinct clusters from their MRG expression profiles. The optimal K value was selected based on various criteria. Based on these criteria, K = 2 was determined to be the most suitable choice (Figs. 4A and S1). Kaplan–Meier analysis indicated that MRG group A patients had a significantly better survival probability (Fig. 4B). The reliability of this classification was further corroborated by results obtained from PCA, tSNE, and UMAP methods, which revealed significant differences among different groups (Fig. 4C). Further analysis indicated significant differences among different groups in clinical and pathological characteristics (Fig. 4D).

Exploration of immunoscape and enrichment analysis

To obtain the biological and immunological roles of the identified MRG clusters, the expression levels of MRGs in each cluster were analyzed. MRG cluster A displayed significantly higher expression of MRGs (Fig. 5A). The level of immune cell infiltration was significantly higher in MRG cluster A (Fig. 5B). In addition, the study employed the GSVA method to identify differential pathways that were enriched in each MRG cluster. MRG cluster A was significantly enriched in pathways associated with intestinal immune network for IGA production, primary immunodeficiency, cell adhesion molecules, asthma, and T cell receptor signaling pathway. On the other hand, MRG cluster B was significantly enriched in pathways associated with drug metabolism, other enzymes, olfactory transduction, steroid biosynthesis, retinol metabolism, and fructose and mannose metabolism (Fig. 5C).
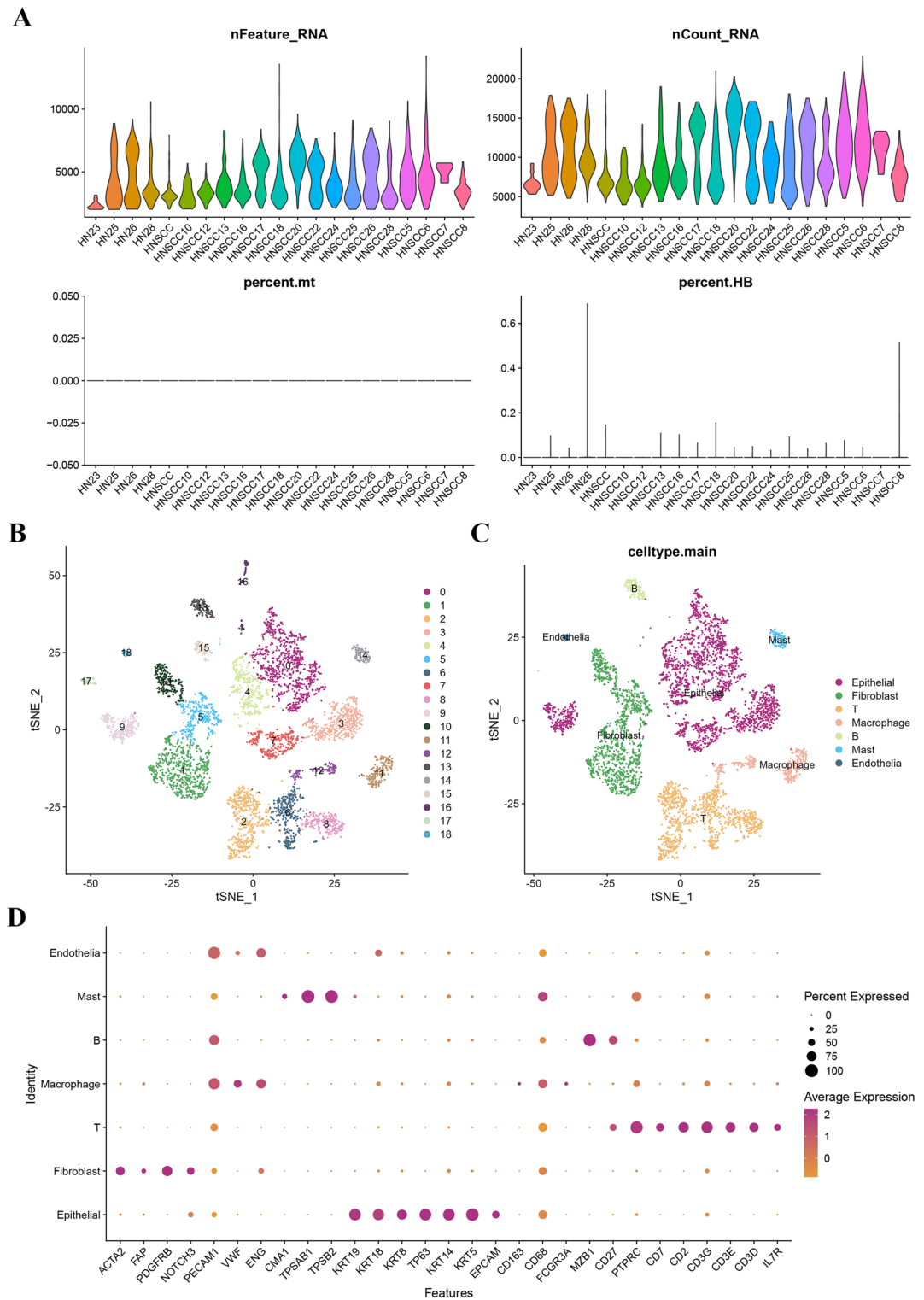
Figure 2. (**A**) 5,902 cells from the GSE103322 database was analyzed after quality control. (**B**) The distribution of different cell populations. (**C**) Cell annotations for each cluster. (**D**) The marker genes.

Construction and validation of MRGs signature

Three sets of HNSCC samples were utilized for analysis, including the training set (n = 358), the testing set (n = 153), and the entire set (n = 511), as well as three external validation sets, namely the GSE27020 set (n = 109), GSE41613 set (n = 97), and GSE65858 set (n = 270). The LASSO analysis was used to recognize a set of 33 prognostic MRGs, as depicted in Fig. 6A. Furthermore, multivariate Cox analysis was applied to construct a prognostic signature consisting of 8 MRGs (Fig. 6B). The tSNE and bubble plots indicated the distribution of expression levels of genes in the model, M1 macrophage markers and M2 macrophage markers (Fig S2). The high-risk
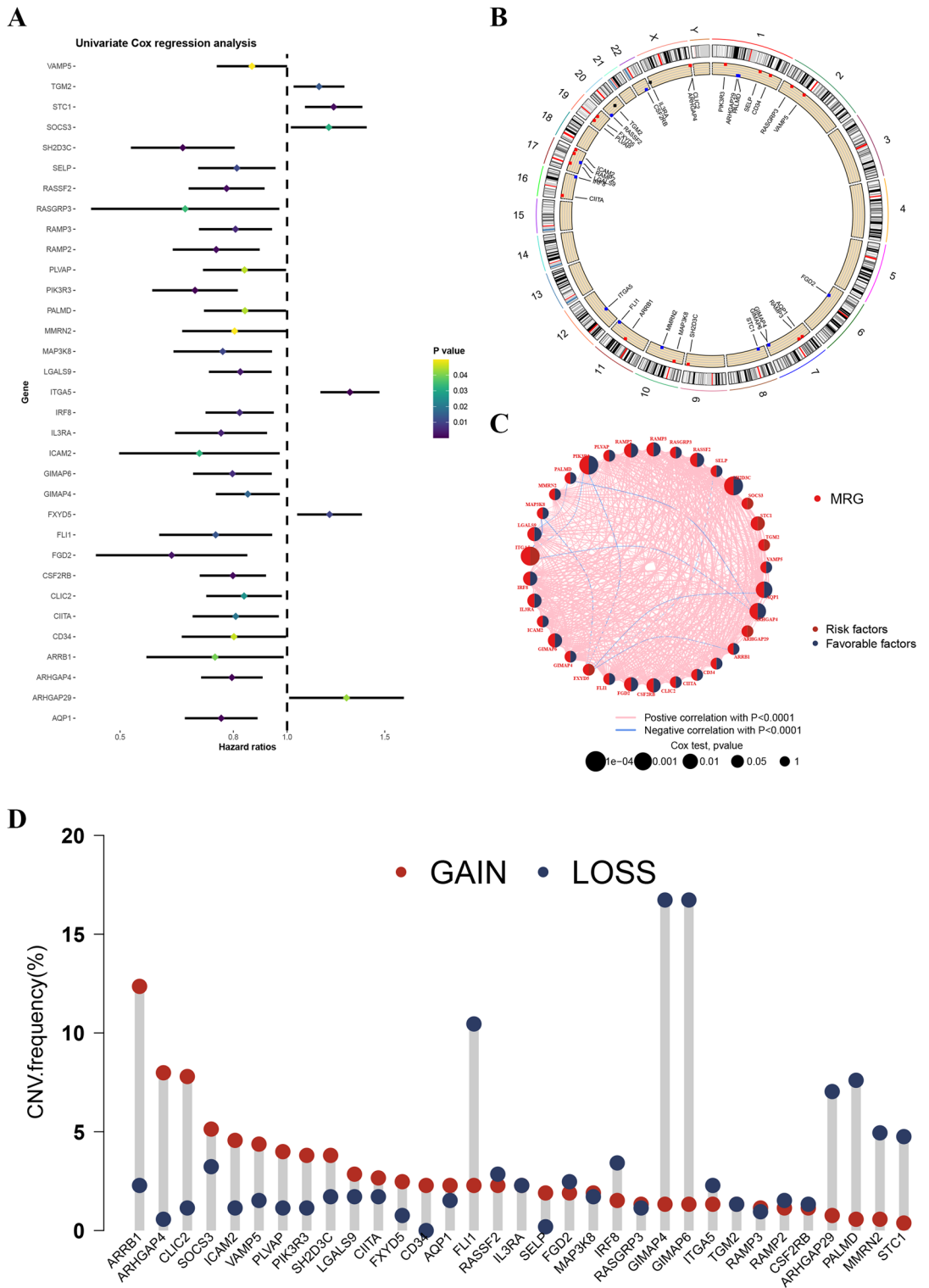
Figure 3. (**A**) 33 prognostic MRGs recognized by univariate Cox analysis. (**B**) The location of 33 MRGs on chromosomes. (**C**) Interaction among 33 MRGs. (**D**) The CNV frequency of 33 MRGs in HNSCC.

patients exhibited shorter survival times (Fig. 6C). This finding was further validated using the external validation sets, wherein the high-risk group consistently showed shorter survival times, confirming the robustness of the MRG signature (Fig. 6C). Additionally, the newly established MRG signature demonstrated the ability to predict 1-, 2-, and 3-year survival rates in HNSCC, with AUC values of 0.688, 0.711, and 0.700 (Fig. 7). Importantly, the AUC values obtained using the MRG signature were higher compared to those obtained using other clinical characteristics, reinforcing the reliability and predictive power of the model (Fig. 7).
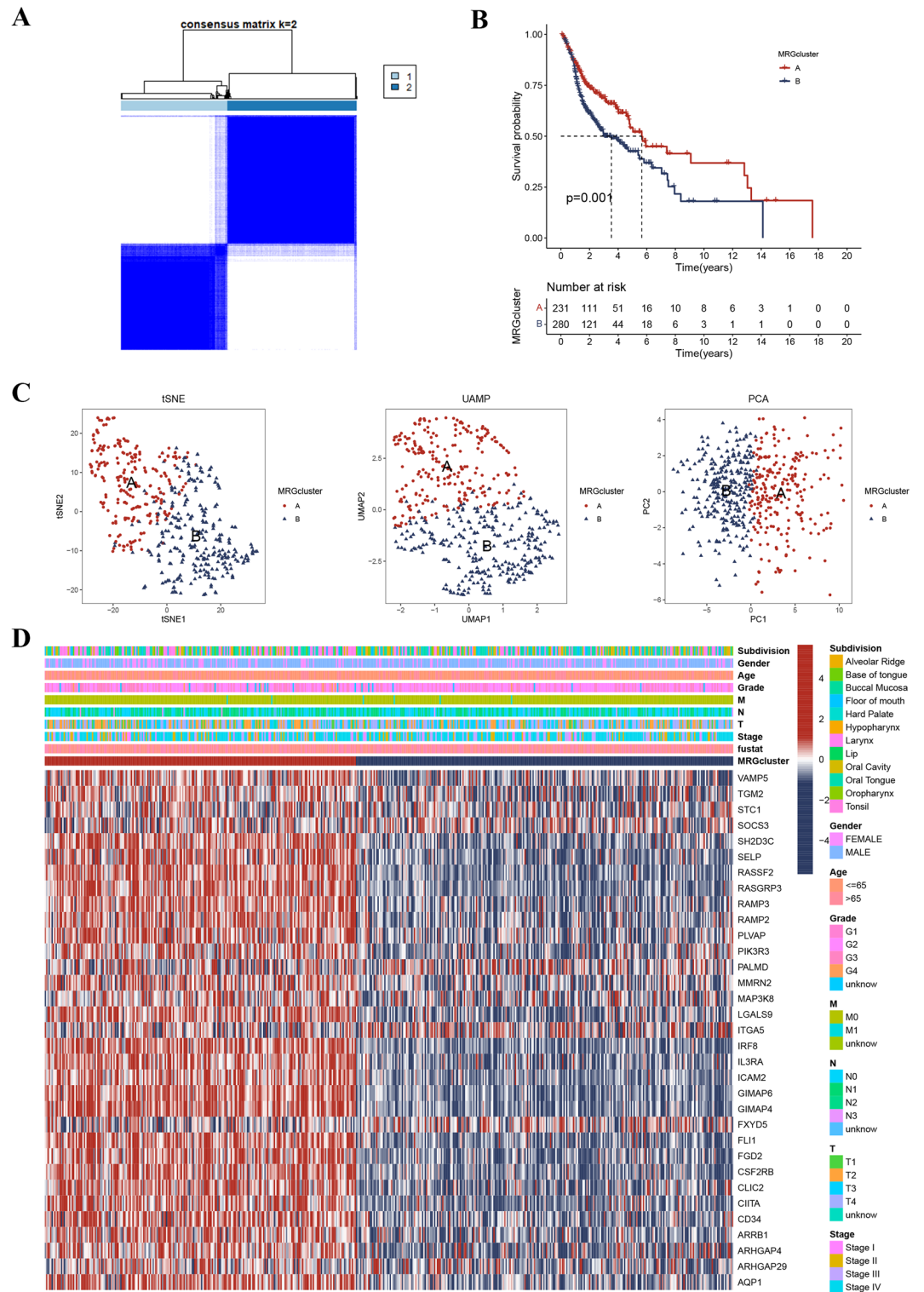
Figure 4. (**A**) Consensus matrix at optimal k = 2. (**B**) Patients in MRG cluster A had a significantly better survival probability. (**C**) The reliability of this classification was further corroborated by results obtained from PCA, tSNE, and UMAP methods. (**D**) Significant differences existed among the two groups in terms of clinical characteristics.

Through analysis of the HPA database, we recognized remarkably higher expression levels of specific proteins, including AQP1, ITGA5, MAP3K8, PIK3R3, STC1, and TGM2, in HNSCC tumor tissues. Conversely, SH2D3C
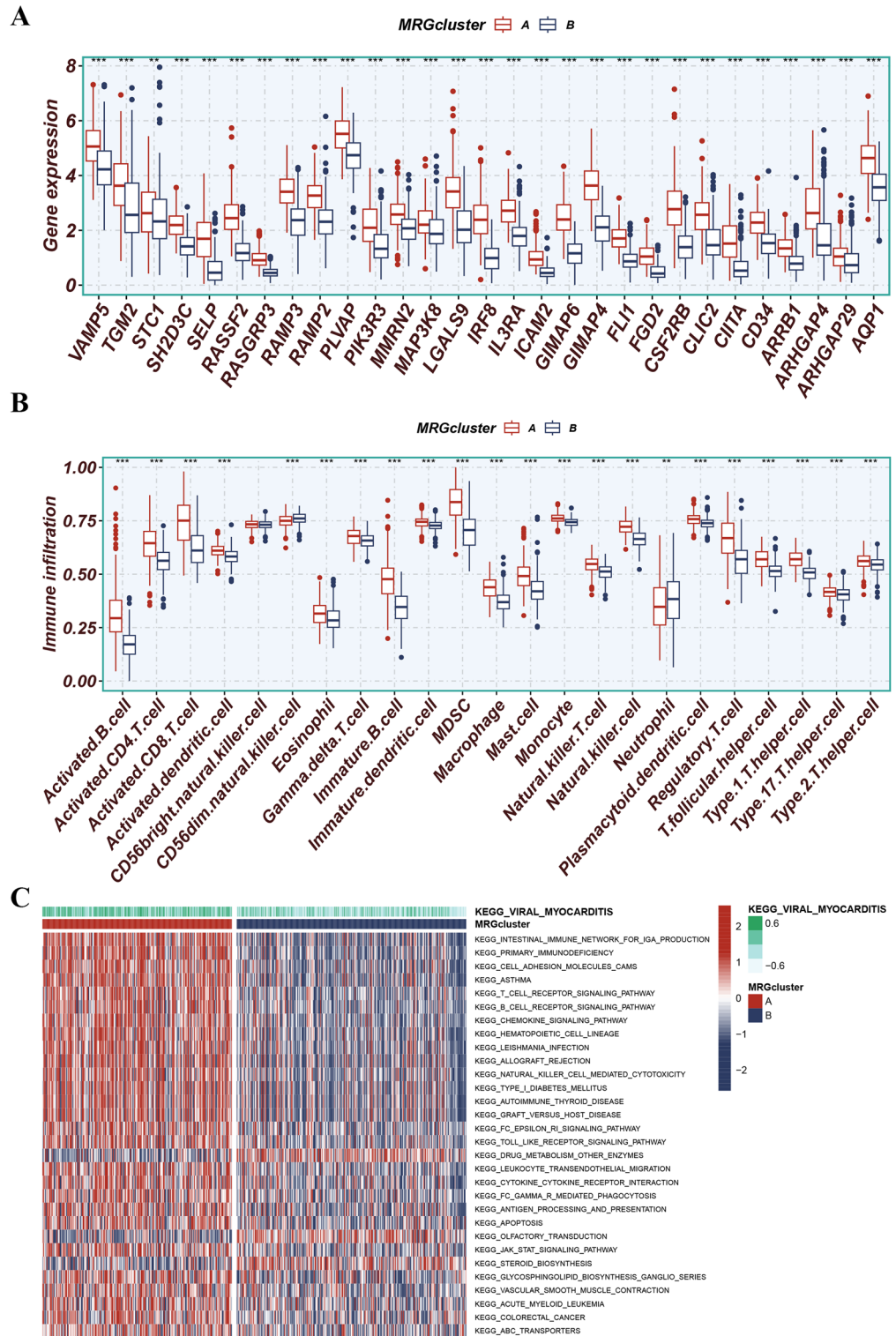
Figure 5. (**A**) MRG cluster A displayed significantly higher expression levels of MRGs. (**B**) The level of immune cell infiltration was significantly higher in MRG cluster A. (**C**) MRG cluster A was significantly enriched in pathways associated with intestinal immune network for IGA production, primary immunodeficiency, cell adhesion molecules CAMs, asthma, and T cell receptor signaling pathway. MRG cluster B was significantly enriched in pathways associated with drug metabolism other enzymes, olfactory transduction, steroid biosynthesis, retinol metabolism, and fructose and mannose metabolism.
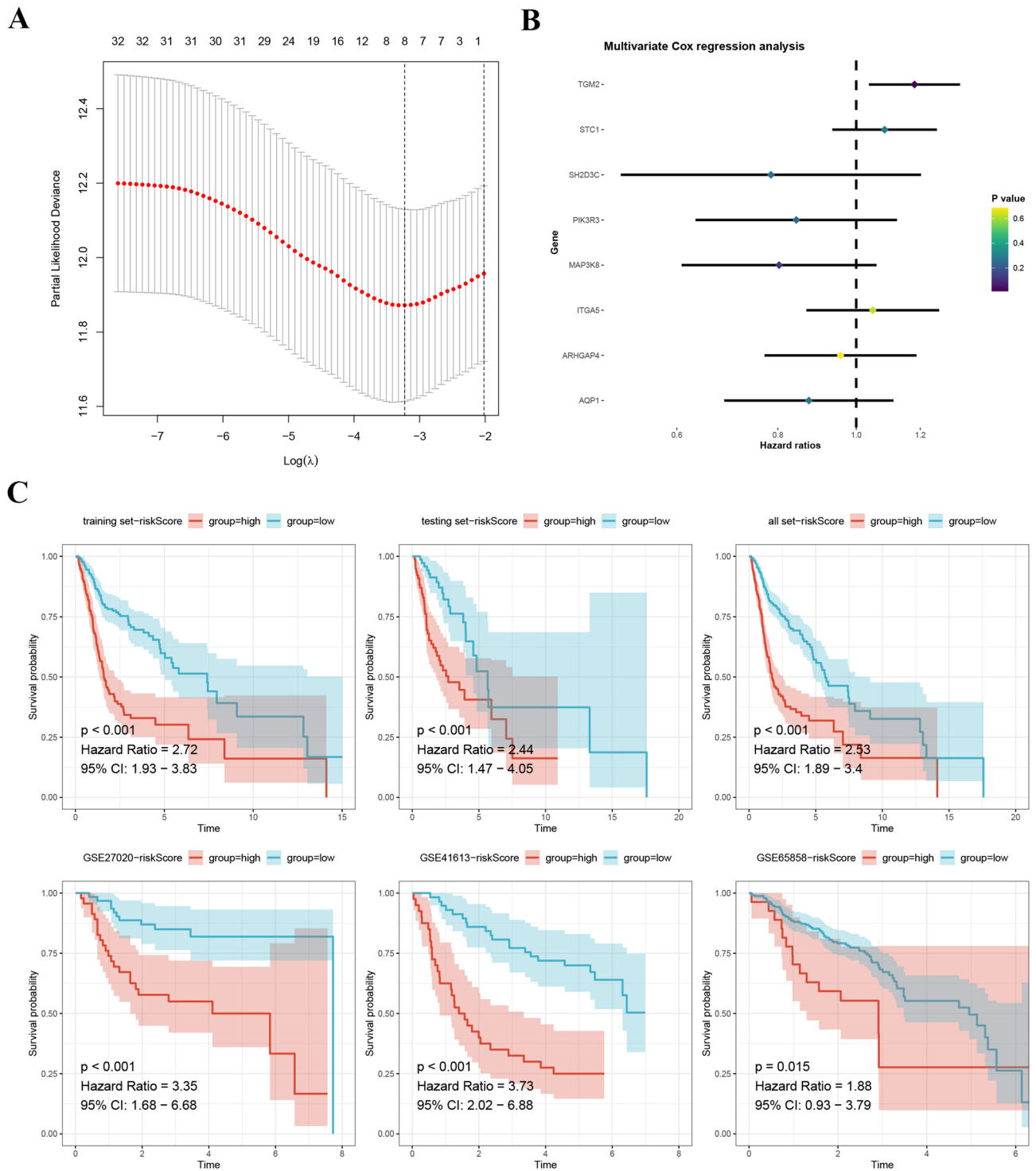
Figure 6. (**A**) 8 prognostic MRGs recognized by LASSO analysis. (**B**) A signature comprising 8 MRGs was developed using multivariate Cox analysis. (**C**) Patients classified as high-risk exhibited significantly shorter survival times.

exhibited remarkably lower expression levels in HNSCC tissues, whereas ARHGAP4 did not show differential expression (Fig. 8). Furthermore, the applicability of the MRG signature to different patient populations was assessed. The lower-risk group patients consistently demonstrated higher survival rates across various clinical subsets, which highlights the usefulness and generalizability of the model in diverse patient populations (Fig. 9). Both univariate and multivariate Cox analyses demonstrated that the risk score was an independent prognostic factor for HNSCC (Fig. 10A and B). Moreover, the C-index further confirmed the superior prognostic performance of the signature compared to clinical features (Fig. 10C). In order to be able to accurately predict the
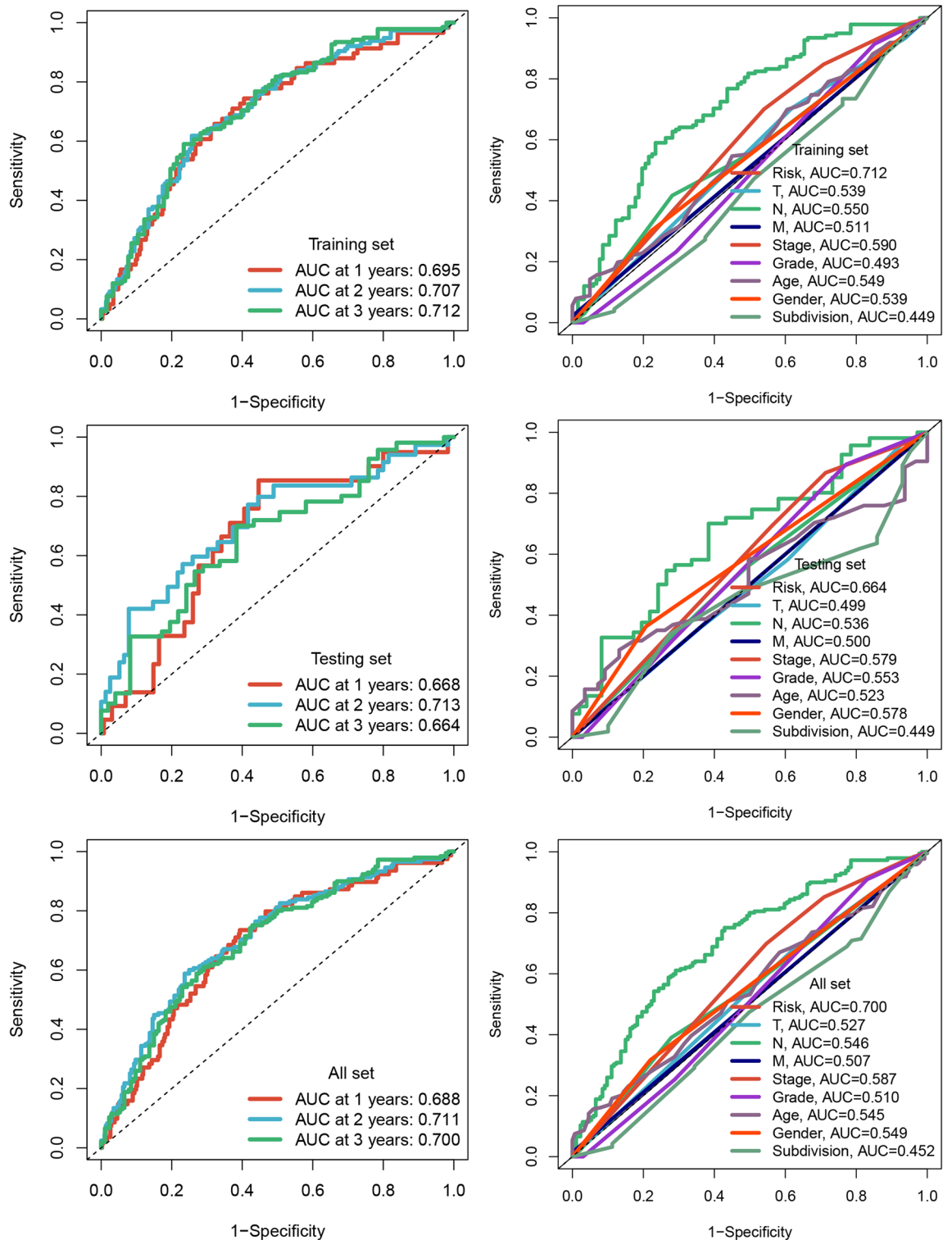
Figure 7. The MRG signature demonstrated the ability to predict survival rates in HNSCC patients. The AUC values obtained using the MRG signature were higher compared to those obtained using other clinical characteristics.

survival of HNSCC, the research developed a nomogram integrating the MRG signature and clinical features (Fig. 10D and E). The nomogram can be used as a useful instrument for clinical decision-making and personalized medicine, providing clinicians with a quantitative and individualized prediction model for HNSCC patient survival.
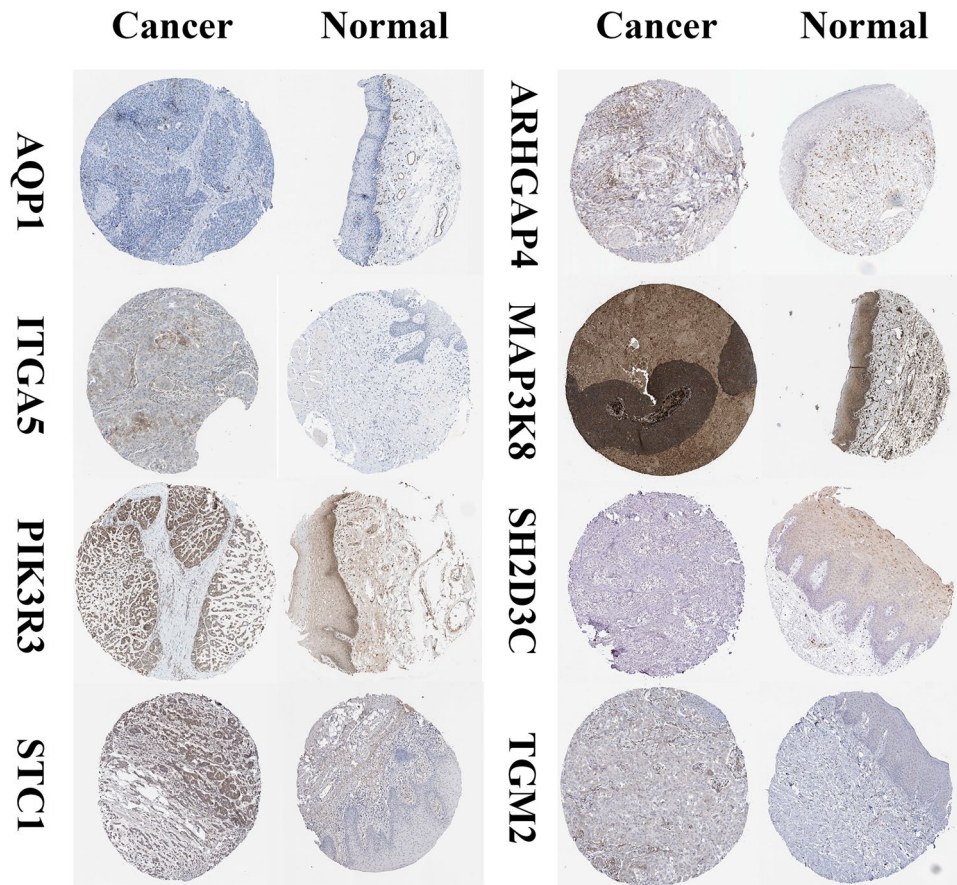
Figure 8. Through analysis of the HPA database, we identified significantly higher expression levels of specific proteins, including AQP1, ITGA5, MAP3K8, PIK3R3, STC1, and TGM2, in HNSCC tumor tissues. Conversely, SH2D3C exhibited significantly lower expression levels in HNSCC tumor tissues, while ARHGAP4 did not show differential expression.

### Enrichment and mutation analysis

132 DEGs were recognized among various groups ($|logFC> = 1.5|$ and FDR < 0.05) (Table S2). BP terms showed enrichment for regulation of mononuclear cell differentiation, lymphocyte differentiation, and B cell activation. CC terms were associated with intermediate filament, intermediate filament cytoskeleton, and the external side of the plasma membrane. MF terms were associated with monooxygenase activity, oxidoreductase activity, and DNA-binding transcription activator activity (Fig. 11A and Table S3). Additionally, KEGG analysis showed that DEGs were highly enriched in pathways including cytokine-cytokine receptor interaction, estrogen signaling pathway, Staphylococcus aureus infection, NF-kappa B signaling pathway, and drug metabolism (Fig. 11B and Table S4). Moreover, the top 10 genes with mutations were identified, and the frequency of mutations was higher in the high-risk group Fig. 11C and D.

Exploration of immunization status

Significant differences in the levels of immune cell subsets including NK cells, B cells, myeloid dendritic cells, cancer-associated fibroblasts, macrophages, mast cells, CD4 + T cells, CD8 + T cells, and regulatory T cells were observed in different groups of immune cell subsets (Fig. 12A). Moreover, differences in immune-related functions such as inflammation-promoting, parainflammation, CCR, checkpoint, T cell co-inhibition, and T cell co-stimulation were observed (Fig. 12B). Furthermore, this study found that the expression levels of several ICGs, including CTLA-4, TIGIT, LAG3, and IDO1, were significantly different between the risk groups (Fig. 13A). Additionally, patients classified into the low-risk group demonstrated lower TIDE scores, indicating a possibility of higher responsiveness to immunotherapy (Fig. 13B). Combining the TIDE scores provided a better prediction of patient prognosis (Fig. 13C). The Sankey plot was informative in revealing the links between MRG clusters, risk groups, and survival status. It showed that exposure to MRG cluster B was related to higher risk scores Fig. 13D and E.

Identification of drugs and correlation with malignant features

70 chemotherapeutic drugs and 81 targeted therapeutic drugs were found, including 5-Fluorouracil, Temozolomide, Carmustine, and EPZ5676, etc. ($p < 0.05$; Fig. 14A and B). Correlation analysis indicated significant
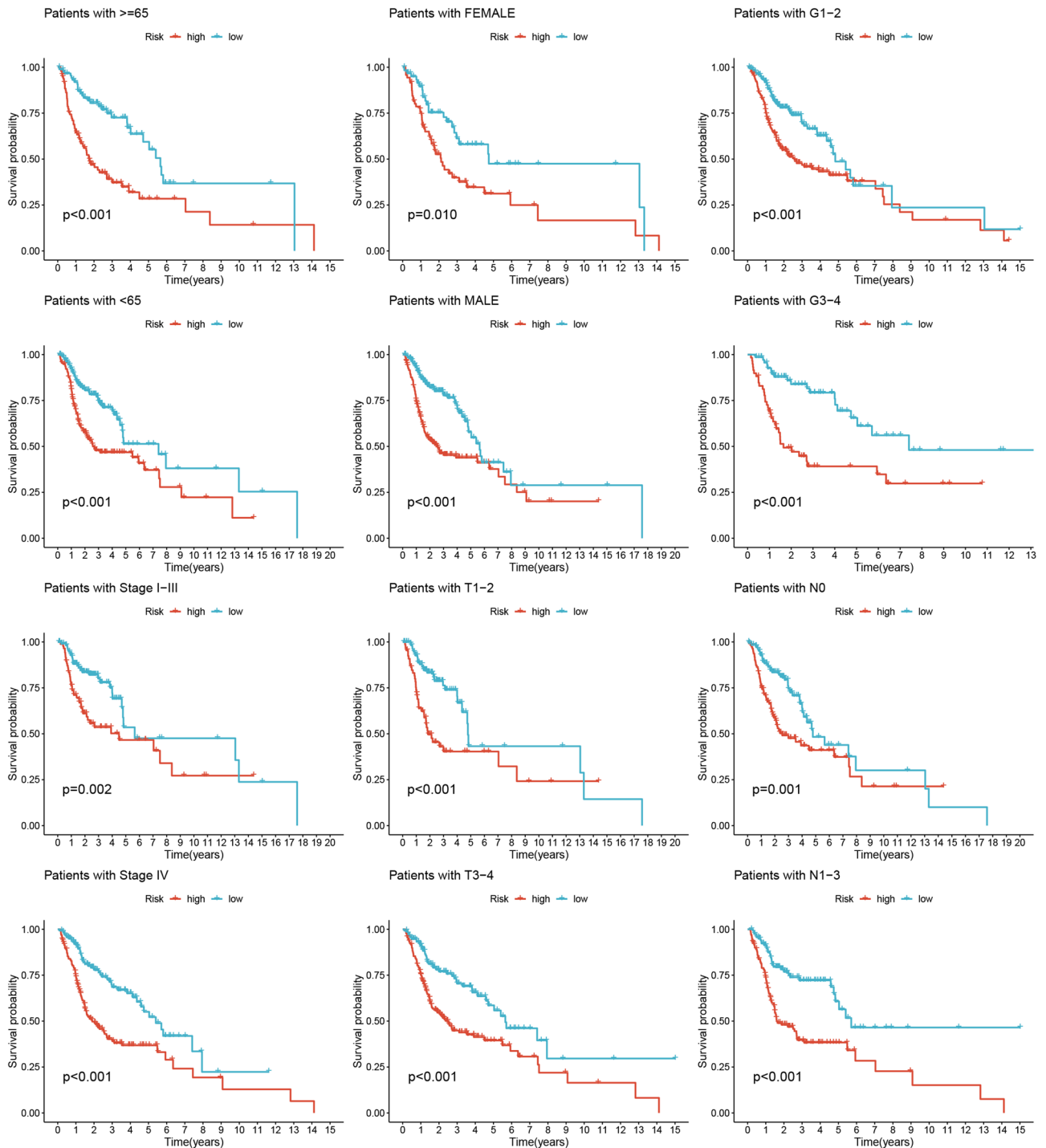
Figure 9. The low-risk group patients indicated higher survival rates across clinical subsets.

correlations between the MRG z-score and angiogenesis z-score (R = 0.74, $p < 0.001$), EMT z-score (R = 0.57, $p < 0.001$), and cell cycle z-score (R = −0.13, $p < 0.001$) across the TCGA pan-cancer cohort (Fig. 15).

Discussion

In the early stage of HNSCC, surgical treatment can often achieve good results, but in the middle and late stages of HNSCC, even if comprehensive treatment such as radiotherapy combined with systemic therapy is used after surgery, the 5-year survival rate is not optimistic, and the incidence of recurrent tumors is as high as 60%[3,55–57]. Immune checkpoint inhibitors have demonstrated good anti-tumor ability by the marker KCSG HN18-12 in patients with HNSCC who have relapsed and advanced metastases after chemotherapy[58]. Li et al.[20] also comprehensively elaborated on the key role of TAM in regulating the progression of head and neck tumors, revealing its immunosuppressive effect and potential mechanism of action in tumor tissues. With the increasing research
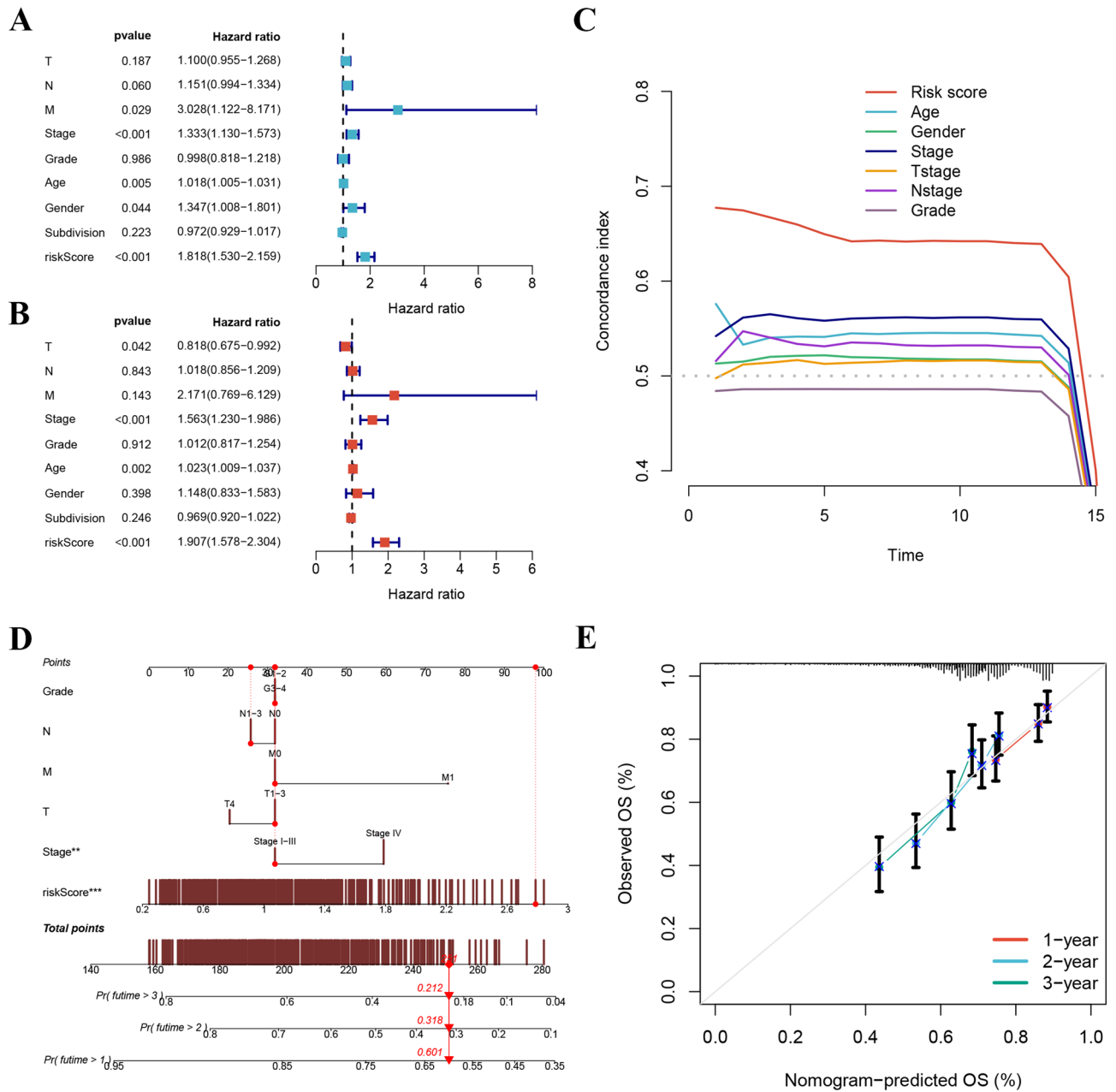
Figure 10. (**A** and **B**) The risk score was an independent prognostic factor. (**C**) C-index confirmed the superior prognostic performance of the signature compared to clinical features. (**D** and **E**) The nomogram demonstrates reliable and sensitive predictions for survival outcomes.

on tumor immunotherapy, treatments including immune checkpoint inhibitors have also brought hope to intermediate and advanced HNSCC patients[59–61].

TAMs are multifunctional tumor-associated immune cells, which often exhibit a variety of different phenotypes according to different tumor types and different TMEs, and the most common phenotypes are M1, M2, etc.[62]. M1 macrophages are considered to be "tumor killer macrophages", secreting pro-inflammatory cytokines, including IL-12, TNF-α, etc., mainly anti-tumor and promote immunity. It has been demonstrated that M1 exosome lncRNA HOTTIP can up-regulate the TLR5/NF-κB pathway by competitive spongy miR-19a-3p and miR-19b-3p, and M1 exosomes and HOTTIP can induce M1 polarization in circulating monocytes, thereby inhibiting the progression of HNSCC[63]. Gene knockout assays have shown that RGS12 can inhibit the progression of oral squamous cell carcinoma by controlling the polarization of M1 in TAMs by controlling ciliated MYCBP2/ KIF2A signaling[64]. The M2 type is the opposite, it secretes anti-inflammatory cytokines and is often thought to be an expression subtype that promotes tumor development.[62,65–67]. Gao et al. used immunohistochemistry and immunofluorescence staining to detect TAM biomarkers and EMT-related proteins, and found that the expression of EMT-related proteins was positively correlated with M2 macrophage biomarkers in HNSCC tissues, revealing that the M2 type of TAMs may induce the EMT process in cancer cells by activating the EGFR/ERK1/2
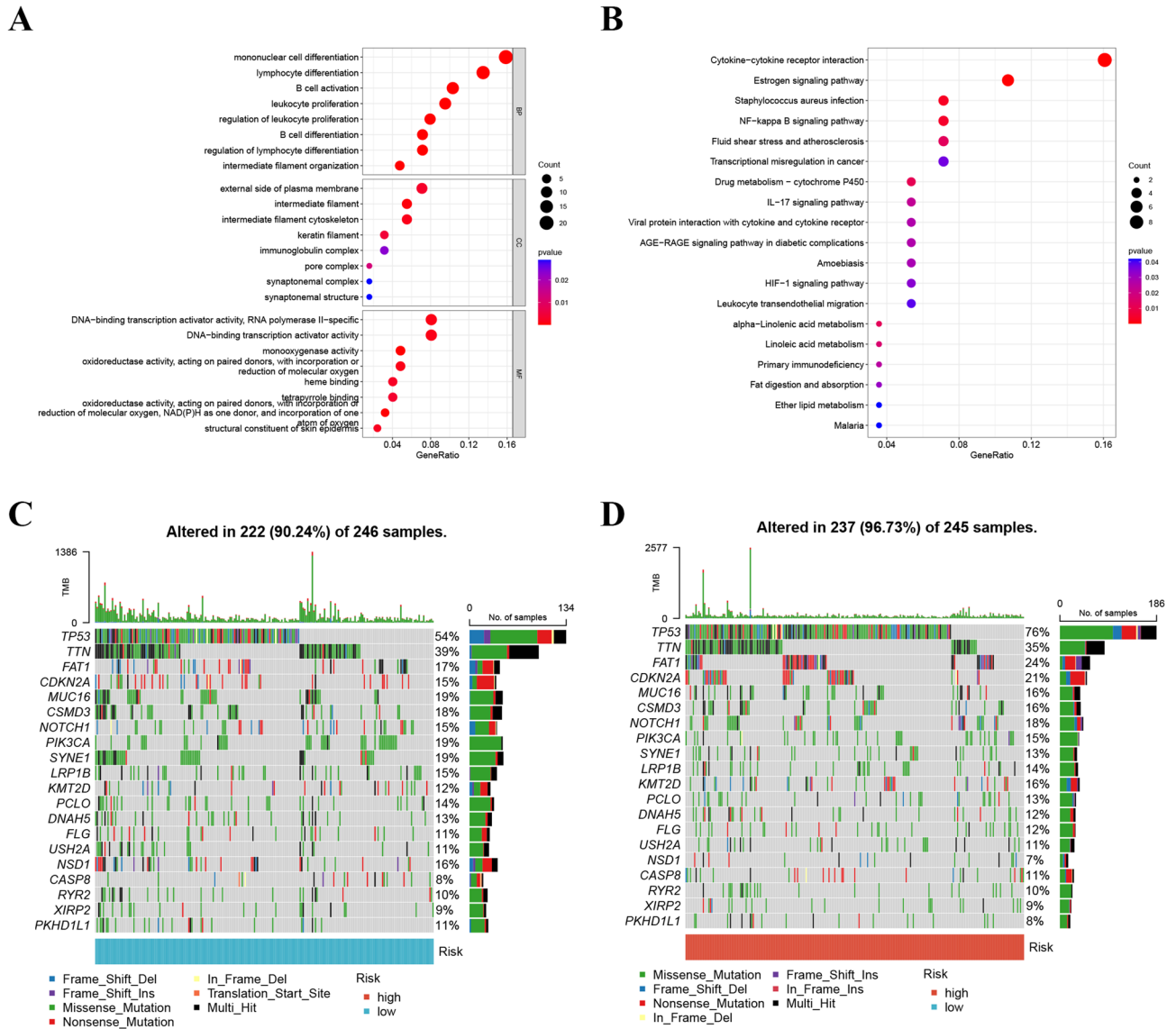
Figure 11. (**A** and **B**) The GO and KEGG analyses for 132 DEGs. (**C** and **D**) The top 10 genes with mutations were identified, and the frequency of mutations was higher in the high-risk group.

signaling pathway in HNSCC and then promote tumors[68]. Through meta-analysis, Ayan et al. showed that the higher density of total TAMs and M2-like subtype TAMs in the TME was associated with T stage progression, nodule positivity, vascular invasion and lymphatic invasion[69]. TAMs are highly related to HNSCC and can help immunotherapy with HNSCC, so MRGs deserve to be explored in depth[70–72]. Therefore, the importance of treating HNSCC accurately and identifying more models helps to control HNSCC, construct a reasonable prognostic risk model, evaluate the correlation with tumor malignant characteristics, explore the immune differences of tumor tissues, and screen out more and more reasonable anti-tumor drugs.

In the study, we first searched for RNA-seq and clinical data related to HNSCC in the open databases GEO-GSE103322, TCGA-HNSCC, GEO-GSE27020, GEO-GSE41613, and GEO-GSE65858. After performing data quality control, we found that 33 MRGs (such as LI1, RASSF2 and IL3RA, etc.) showed the potential for disease prognosis prediction and were significantly related to survival, of which 18 genes have potential carcinogenic effects, and 15 other genes have potential cancer suppression. For these genes, some researchers have found that ARRB1 can promote the activation of the TAK1/MAPK pathway to promote the progression of gallbladder cancer, and ARHGAP4 can be targeted by miR-939-5p, thereby promoting the invasion and metastasis of pancreatic cancer[73,74]. Experiments have shown that FXYD5 can promote metastasis of mouse breast cancer tissues by regulating the β-Na + -K + -ATPase subunit[75]. Similarly, in another type of gene, the gene RASSF2 was shown to be an oncogene for Ewing sarcoma, and high GIMAP4 expression could facilitate the TME of immune cells to detect tumor cells and inhibit the development of lung adenocarcinoma cells[76,77]. There are many similar scientific studies mentioned above, and the effects of these genes on tumors in the above studies are consistent and trustworthy with the conclusions of this study. The consensus clustering method divided patients into different clusters, and the survival probability analysis showed that the survival rate of group A patients was much higher.
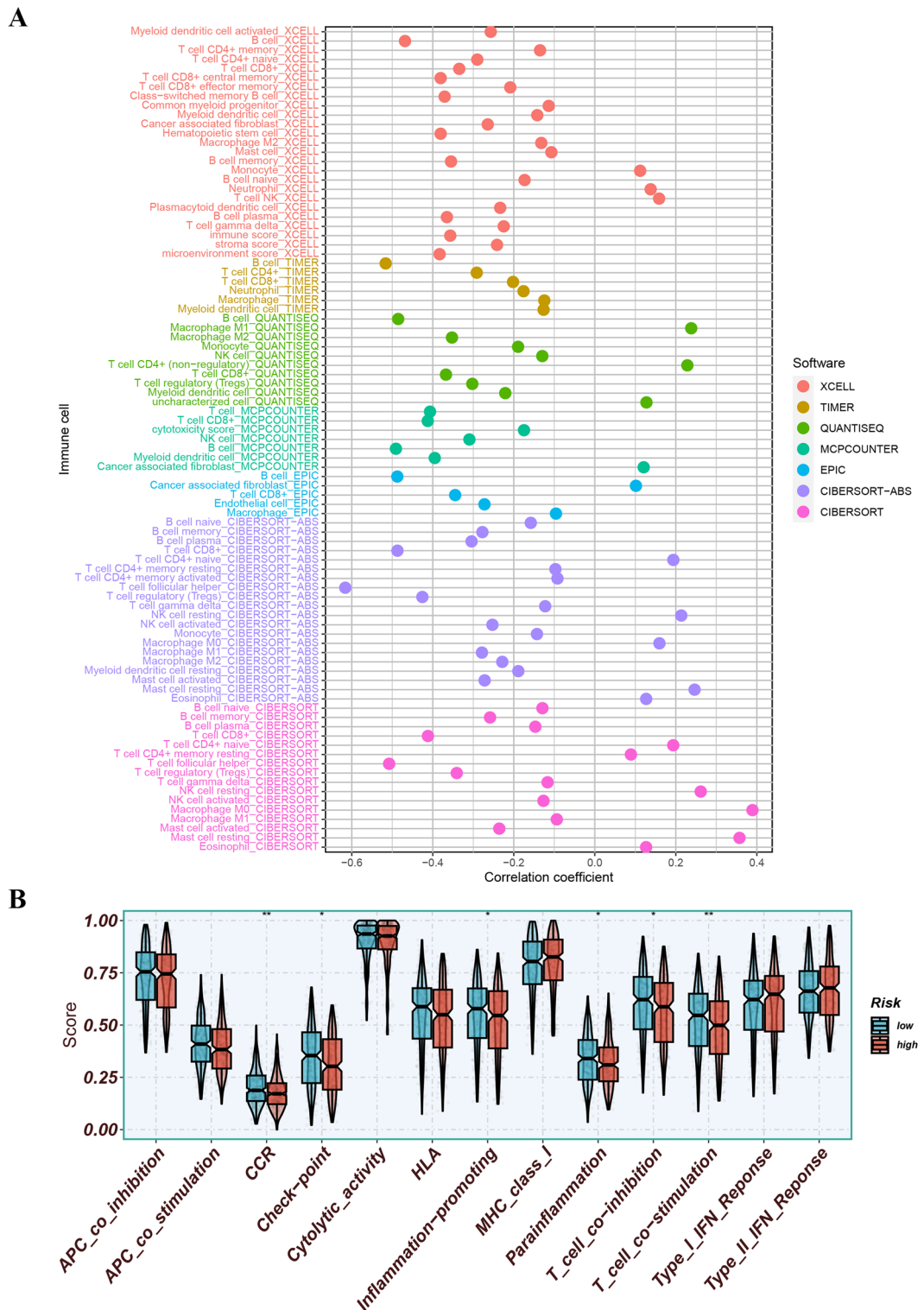
Figure 12. (**A**) Differences in immune cell subsets. (**B**) Differences in immune-related functions.

We found that they had different features by analyzing differences in clinical features among different subtypes, suggesting that this might be a useful prognostic marker for patients with HNSCC, which will likely guide our clinical work. Gene expression analysis showed that cluster A was more closely correlated with MRG, and GSEA analysis found that the immune cell infiltration level of cluster A was significantly increased. GSVA analysis found that cluster A was significantly enriched in pathways related to IGA production and primary immunodeficiency, while cluster B was more enriched in drug metabolism and olfactory transduction of other enzymes. In the face of different enrichment states, this will be able to guide us to differentiated treatment of patients.
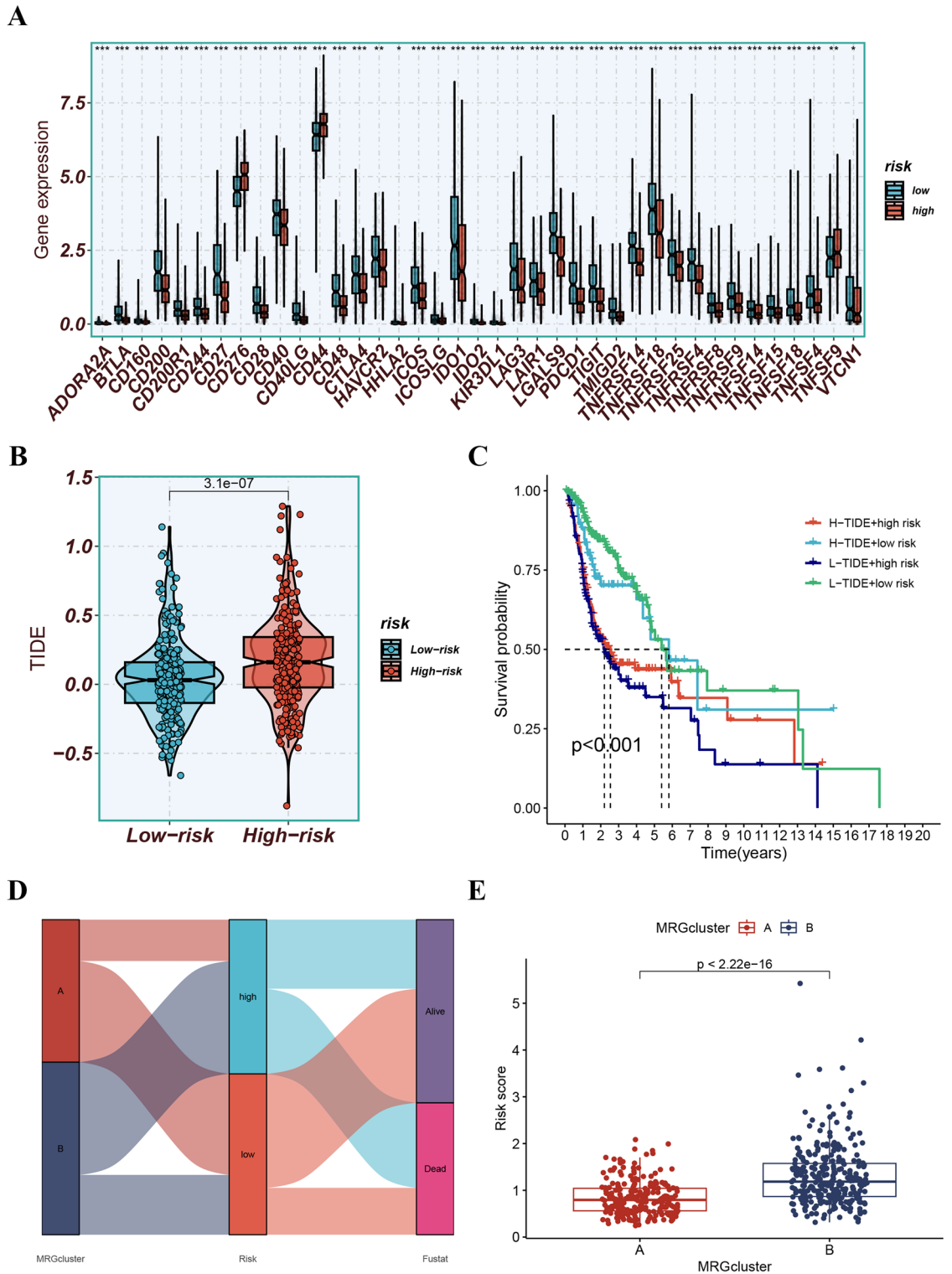
Figure 13. (**A**) The expression levels of various ICGs, including CTLA-4, TIGIT, LAG3, and IDO1, were significantly different between the risk groups. (**B**) Patients classified into the low-risk group demonstrated lower TIDE scores, indicating a possibility of higher responsiveness to immunotherapy. (**C**) Combining the TIDE scores provided a better prediction of patient prognosis. (**D** and **E**) The Sankey plot was informative in revealing the links between MRG clusters, risk groups, and survival status. It showed that exposure to MRG cluster B was associated with higher risk scores.
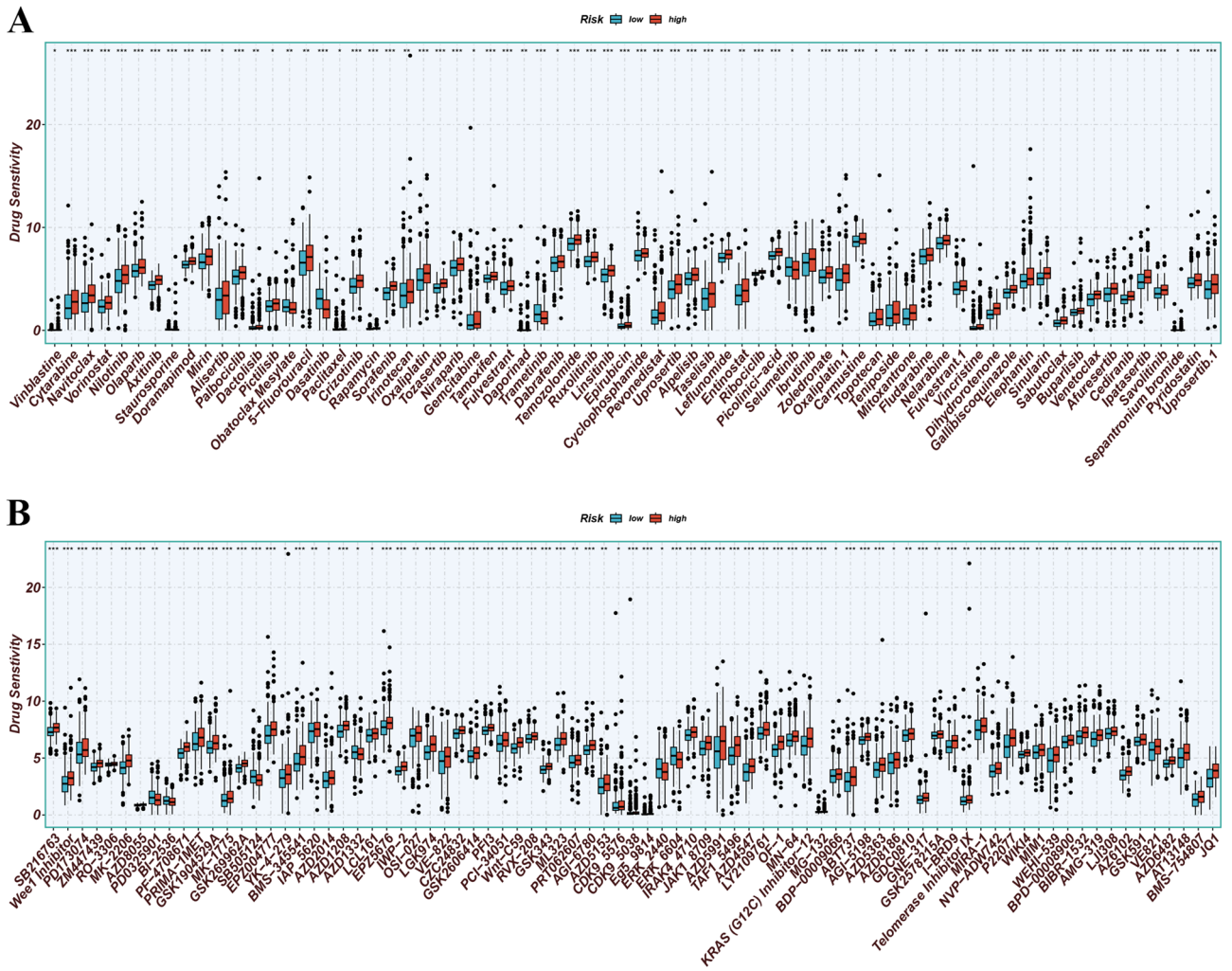
Figure 14. (**A** and **B**) A search for chemotherapeutic medications and targeted agents was conducted, and customized therapy regimens were developed based on patient subgroups ($p < 0.05$).

In this study, 3 sets of HNSCC samples were used for analysis, 33 prognostic MRGs were identified by LASSO analysis, and a prognostic model consisting of 8 MRGs (TGM2, STC1, SH2D3C, PIK3R3, MAP3K8, ITGA5, ARHGAP4 and AQP1) was constructed by multivariate Cox analysis. The survival prognosis was notably worse in the high-risk group, and this finding was further validated by the external validation set. The newly established MRG model can better forecast the survival rates of HNSCC, and its higher AUC value enhances the reliability of the model, which will guide whether clinical treatment is aggressive or palliative. In this study, AQP1, ITGA5, MAP3K8, PIK3R3, STC1 and TGM2 were significantly overexpressed in HNSCC tumor tissues, while SH2D3C expression was significantly reduced, and there was no difference in ARHGAP4 expression. This is similar to the results of "Guo et al. found that AQP1 can promote local invasion of breast cancer, Xu et al. found that ITGA5 can promote angiogenesis in cervical cancer, and Lee et al. found that MAP3K8 overexpression can induce squamous cell carcinogenesis in the salivary glands of mice"[78–80]. Studies have shown that STC1 is not only one of MRGs, but also a glycolysis-related genes in HNSCC, which is prominently expressed in the glycolytic activities of tumor tissues and can also predict the prognosis of HNSCC from the direction of glycolysis, which is similar to the conclusion of this study[81]. In this study, we also integrated MRG features and clinical features to develop a nomogram, which will provide an excellent tool for clinical and scientific work.

At the same time, we identified 132 DEGs between different groups, and GO analysis and KEGG analysis further indicated the different roles of DEGs. In addition, we analyzed mutations in the first 10 typical genes, whose mutations might facilitate the progression of HNSCC. Among them, TP53 and other genes are hot genes in the research of HNSCC, which has great immunotherapy prospects[82–85]. Studies have confirmed that FAT1 mutations are highly abundant in cisplatin responders in HNSCC patients, and the gene has been found to have potential targetable changes in 15% of HNSCC patients, which is great news for targeted immunotherapy[86]. The immune landscape study found significant differences in multiple immune cell subsets and in several ICG groups, and the low-risk group showed a lower TIDE score, suggesting a possible higher response to immunotherapy. Therefore, pooled TIDE scores provide a better predictor of patient outcomes.
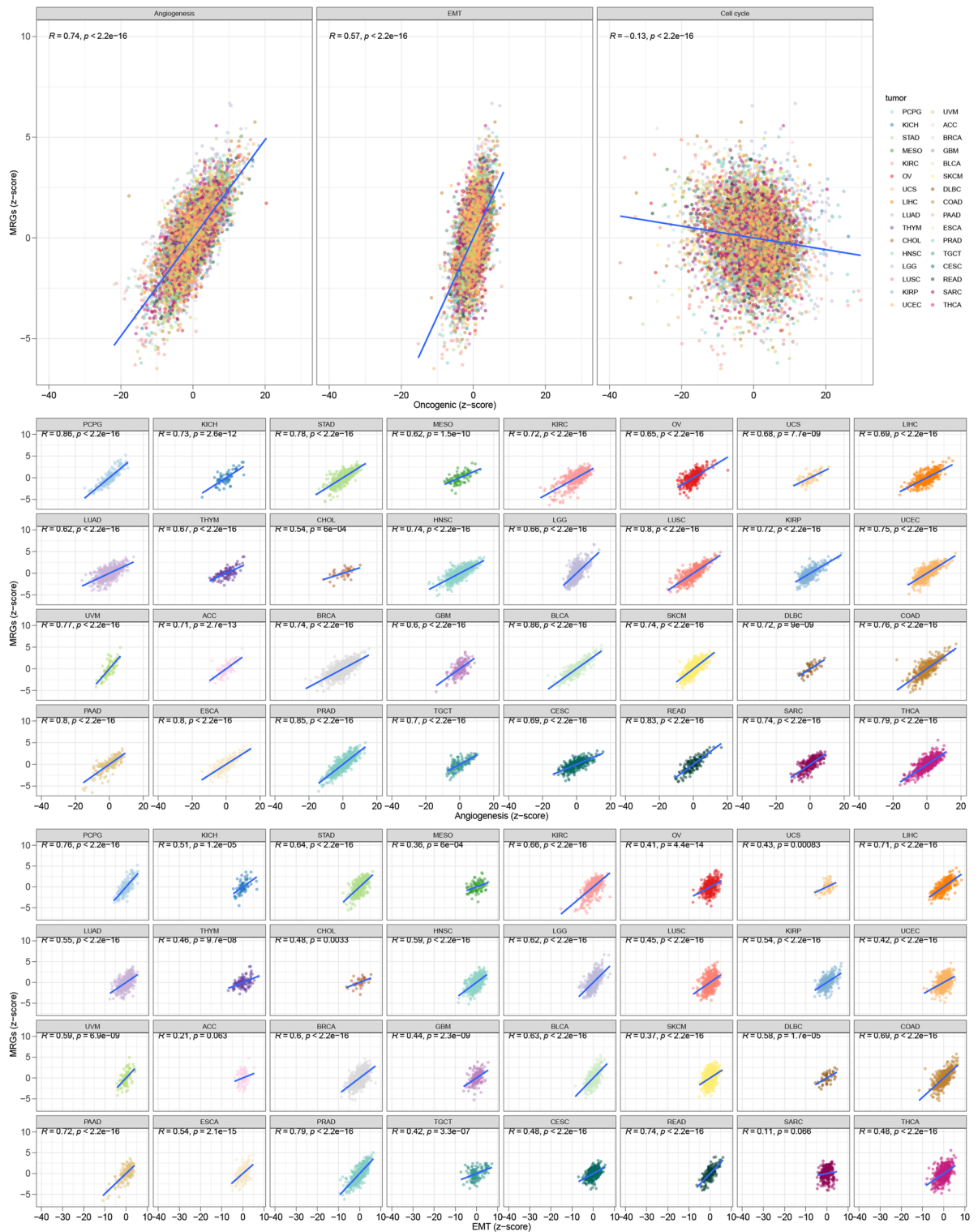
Figure 15. Significant correlations between the MRG z-score and angiogenesis z-score (R = 0.74, *p* < 0.001), EMT z-score (R = 0.57, *p* < 0.001), and cell cycle z-score (R = -0.13, *p* < 0.001).

Through the search for various chemotherapy drugs and innovative targeted drugs, this may lead to the generation of personalized treatment regimens for patients belonging to different subgroups. The more prominent drugs are 5-Fluorouracil, Temozolomide, Carmustine, and EPZ5676, etc. At the end of the study, using the z-scoring algorithm, we found significant associations with angiogenesis, EMT, and cell cycle in the TCGA cancer cohort.

Even so, our research had certain limitations. Firstly, the research is based on an analysis of public databases, which is a retrospective study. Even if scientific verification methods are used in the study, the absolute accuracy

and applicability of the relevant conclusions cannot be guaranteed, and a large amount of experimental data and clinical data are still required to verify the relevant models and conclusions.

Conclusion

We constructed a prognostic signature consisting of 8 MRGs (TGM2, STC1, SH2D3C, PIK3R3, MAP3K8, ITGA5, ARHGAP4 and AQP1) by multivariate Cox analysis. The low-risk group of patients may have a higher response to immunotherapy. In terms of drug selection, the more prominent drugs are 5-Fluorouracil, Temozolomide and so on. Angiogenesis, EMT, and cell cycle are malignant features associated with HNSCC. This study opens up new prospects for the prognosis prediction and clinical treatment strategy of HNSCC.

Data availability

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* **71**, 209–249 (2021).
2. Yin, J. *et al.* Identification of molecular classification and gene signature for predicting prognosis and immunotherapy response in HNSCC using cell differentiation trajectories. *Sci Rep.* **12**, 20404 (2022).
3. Mody, M. D., Rocco, J. W., Yom, S. S., Haddad, R. I. & Saba, N. F. Head and neck cancer. *Lancet.* **398**, 2289–2299 (2021).
4. Agalliu, I. *et al.* Associations of oral alpha-, beta-, and gamma-human papillomavirus types with risk of incident head and neck cancer. *JAMA Oncol.* **2**, 599–606 (2016).
5. Fakhry, C. & Gillison, M. L. Clinical implications of human papillomavirus in head and neck cancers. *J. Clin. Oncol.* **24**, 2606–2611 (2006).
6. Chang, J. *et al.* Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. *Nat. Commun.* **8**, 15290 (2017).
7. Yokota, T. *et al.* Immunotherapy for squamous cell carcinoma of the head and neck. *Jpn. J. Clin. Oncol.* **50**, 1089–1096 (2020).
8. Maghami, E. *et al.* Diagnosis and management of squamous cell carcinoma of unknown primary in the head and neck: ASCO guideline. *J. Clin. Oncol.* **38**, 2570–2596 (2020).
9. Koyfman, S. A. *et al.* Management of the neck in squamous cell carcinoma of the oral cavity and oropharynx: ASCO clinical practice guideline. *J. Clin. Oncol.* **37**, 1753–1774 (2019).
10. Ruffin, A. T. *et al.* Improving head and neck cancer therapies by immunomodulation of the tumour microenvironment. *Nat. Rev. Cancer.* **23**, 173–188 (2023).
11. Moskovitz, J. M., Moy, J., Seiwert, T. Y. & Ferris, R. L. Immunotherapy for head and neck squamous cell carcinoma: A review of current and emerging therapeutic options. *Oncologist* **22**, 680–693 (2017).
12. Subramaniam, S. S., Paterson, C. & McCaul, J. A. Immunotherapy in the management of squamous cell carcinoma of the head and neck. *Br J Oral Maxillofac Surg.* **57**, 957–966 (2019).
13. Nocini, R. *et al.* ICGC ARGO precision medicine: Genomic profiling-informed prediction of immunotherapy response in two patients with metastatic head and neck squamous cell carcinoma. *Lancet Oncol.* **24**, 851–853 (2023).
14. Mantovani, A., Allavena, P., Marchesi, F. & Garlanda, C. Macrophages as tools and targets in cancer therapy. *Nat. Rev. Drug Discov.* **21**, 799–820 (2022).
15. Zhang, X. *et al.* Tumor-associated M2 macrophages in the immune microenvironment influence the progression of renal clear cell carcinoma by regulating M2 macrophage-associated genes. *Front Oncol.* **13**, 1157861 (2023).
16. Attri, K. S., Mehla, K., Shukla, S. K. & Singh, P. K. Microscale gene expression analysis of tumor-associated macrophages. *Sci Rep.* **8**, 2408 (2018).
17. Wu, L. *et al.* Exosomes derived from gastric cancer cells activate NF-kappaB pathway in macrophages to promote cancer progression. *Tumour Biol.* **37**, 12169–12180 (2016).
18. Sun, Q. Y. *et al.* Analysis between macrophage-related genes with prognosis and tumor microenvironment in non-small cell lung cancer. *Yi Chuan.* **45**, 684–699 (2023).
19. Xu, B. *et al.* Establishment and validation of a genetic label associated with M2 macrophage infiltration to predict survival in patients with colon cancer and to assist in immunotherapy. *Front Genet.* **12**, 726387 (2021).
20. Li, B., Ren, M., Zhou, X., Han, Q. & Cheng, L. Targeting tumor-associated macrophages in head and neck squamous cell carcinoma. *Oral Oncol.* **106**, 104723 (2020).
21. Puram, S. V. *et al.* Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell.* **171**, 1611–1624 (2017).
22. Fountzilas, E. *et al.* Identification and validation of a multigene predictor of recurrence in primary laryngeal cancer. *PLoS One.* **8**, e70429 (2013).
23. Lohavanichbutr, P. *et al.* A 13-gene signature prognostic of HPV-negative OSCC: Discovery and external validation. *Clin. Cancer Res.* **19**, 1197–1203 (2013).
24. Zhao, Y. *et al.* Comprehensive analysis of tumor immune microenvironment characteristics for the prognostic prediction and immunotherapy of oral squamous cell carcinoma. *Front Genet.* **13**, 788580 (2022).
25. Wichmann, G. *et al.* The role of HPV RNA transcription, immune response-related gene expression and disruptive TP53 mutations in diagnostic and prognostic profiling of head and neck cancer. *Int. J. Cancer.* **137**, 2846–2857 (2015).
26. Zhang, Y. *et al.* A signature for pan-cancer prognosis based on neutrophil extracellular traps. *J. Immunother. Cancer.* **10**, e004210 (2022).
27. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
28. Chen, S. *et al.* Single-cell analysis reveals transcriptomic remodellings in distinct cell types that contribute to human prostate cancer progression. *Nat. Cell Biol.* **23**, 87–98 (2021).
29. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
30. Zhang, H., Meltzer, P. & Davis, S. RCircos: An R package for circos 2D track plots. *BMC Bioinform.* **14**, 244 (2013).
31. Zhang, Z. Reshaping and aggregating data: An introduction to reshape package. *Ann. Transl. Med.* **4**, 78 (2016).
32. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).

33. Yin, J. *et al.* Head and neck squamous cell carcinoma subtypes based on immunologic and hallmark gene sets in tumor and non-tumor tissues. *Front. Surg.* **9**, 821600 (2022).

34. Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **14**, 7 (2013).

35. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).

36. Simon, N., Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* **39**, 1–13 (2011).

37. Blanche, P., Dartigues, J. F. & Jacqmin-Gadda, H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat. Med.* **32**, 5381–5397 (2013).

38. Wu, T. *et al.* ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation.* **2**, 100141 (2021).

39. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

40. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

41. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

42. Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C. & Koeffler, H. P. Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **28**, 1747–1756 (2018).

43. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods.* **12**, 453–457 (2015).

44. Li, T. *et al.* TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* **77**, e108–e110 (2017).

45. Dienstmann, R. *et al.* Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. *Ann. Oncol.* **30**, 1622–1629 (2019).

46. Aran, D., Hu, Z. & Butte, A. J. XCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **18**, 220 (2017).

47. Finotello, F. *et al.* Correction to: Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med.* **11**, 50 (2019).

48. Tamminga, M. *et al.* Immune microenvironment composition in non-small cell lung cancer and its association with survival. *Clin Transl. Immunol.* **9**, e1142 (2020).

49. Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E. & Gfeller, D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *ELife* https://doi.org/10.7554/eLife.26476 (2017).

50. Fu, J. *et al.* Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* **12**, 21 (2020).

51. Maeser, D., Gruener, R. F. & Huang, R. S. OncoPredict: An R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Brief. Bioinform.* https://doi.org/10.1093/bib/bbab260 (2021).

52. Lee, E., Chuang, H. Y., Kim, J. W., Ideker, T. & Lee, D. Inferring pathway activity toward precise disease classification. *PLoS Comput. Biol.* **4**, e1000217 (2008).

53. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-Cell RNA-seq. *Science* **352**, 189–196 (2016).

54. Senbabaoglu, Y. *et al.* Erratum to: Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* **18**, 46 (2017).

55. Johnson, D. E. *et al.* Author correction: Head and neck squamous cell carcinoma. *Nat Rev Dis Primers.* **9**, 4 (2023).

56. Cramer, J. D., Burtness, B., Le, Q. T. & Ferris, R. L. The changing therapeutic landscape of head and neck cancer. *Nat. Rev. Clin. Oncol.* **16**, 669–683 (2019).

57. Watermann, C. *et al.* Recurrent HNSCC harbor an immunosuppressive tumor immune microenvironment suggesting successful tumor immune evasion. *Clin. Cancer Res.* **27**, 632–644 (2021).

58. Lee, Y. G. *et al.* Outcomes and biomarkers of immune checkpoint inhibitor therapy in patients with refractory head and neck squamous cell carcinoma: KCSG HN18-12. *Cancer Res. Treat.* **53**, 671–677 (2021).

59. Carlisle, J. W., Steuer, C. E., Owonikoko, T. K. & Saba, N. F. An update on the immune landscape in lung and head and neck cancers. *CA Cancer J. Clin.* **70**, 505–517 (2020).

60. Sacco, A. G. *et al.* Pembrolizumab plus cetuximab in patients with recurrent or metastatic head and neck squamous cell carcinoma: An open-label, multi-arm, non-randomised, multicentre, phase 2 trial. *Lancet Oncol.* **22**, 883–892 (2021).

61. Ho, A. L. Immunotherapy, chemotherapy, or both: Options for first-line therapy for patients with recurrent or metastatic head and neck squamous cell carcinoma. *J. Clin. Oncol.* **41**, 736–741 (2023).

62. Muri, J., Chylikova, J., Skarda, J., Miklosova, M. & Kamarad, V. The role of tumor-associated macrophages in solid malignancies—An overview of current knowledge. *Biomed. Pap. Med. Fac. Univ. Palacky Olomouc. Czech. Repub.* **166**, 136–139 (2022).

63. Jiang, H. *et al.* M1 macrophage-derived exosomes and their key molecule lncRNA HOTTIP suppress head and neck squamous cell carcinoma progression by upregulating the TLR5/NF-kappaB pathway. *Cell Death Dis.* **13**, 183 (2022).

64. Yuan, G., Yang, S. & Yang, S. RGS12 represses oral squamous cell carcinoma by driving M1 polarization of tumor-associated macrophages via controlling ciliary MYCBP2/KIF2A signaling. *Int. J. Oral Sci.* **15**, 11 (2023).

65. Hu, W. *et al.* Tumor-associated macrophages in cancers. *Clin. Transl. Oncol.* **18**, 251–258 (2016).

66. Shu, Y. & Cheng, P. Targeting tumor-associated macrophages for cancer immunotherapy. *Biochim. Biophys. Acta Rev. Cancer.* **1874**, 188434 (2020).

67. Pan, Y., Yu, Y., Wang, X. & Zhang, T. Tumor-associated macrophages in tumor immunity. *Front. Immunol.* **11**, 583084 (2020).

68. Gao, L. *et al.* Tumor associated macrophages induce epithelial to mesenchymal transition via the EGFR/ERK1/2 pathway in head and neck squamous cell carcinoma. *Oncol. Rep.* **40**, 2558–2572 (2018).

69. Kumar, A. T. *et al.* Prognostic significance of tumor-associated macrophage content in head and neck squamous cell carcinoma: a meta-analysis. *Front. Oncol.* **9**, 656 (2019).

70. Lu, T. *et al.* CD73 in small extracellular vesicles derived from HNSCC defines tumour-associated immunosuppression mediated by macrophages in the microenvironment. *J. Extracell. Vesicles* **11**, e12218 (2022).

71. Affara, N. I. *et al.* B cells regulate macrophage phenotype and response to chemotherapy in squamous carcinomas. *Cancer Cell.* **25**, 809–821 (2014).

72. Ludwig, N. *et al.* TGFbeta(+) small extracellular vesicles from head and neck squamous cell carcinoma cells reprogram macrophages towards a pro-angiogenic phenotype. *J. Extracell. Vesicles.* **11**, e12294 (2022).

73. Zhang, X. *et al.* ARRB1 drives gallbladder cancer progression by facilitating TAK1/MAPK signaling activation. *J. Cancer.* **12**, 1926–1935 (2021).

74. Shen, Y. *et al.* MiR-939-5p contributes to the migration and invasion of pancreatic cancer by targeting ARHGAP4. *Onco Targets Ther.* **13**, 389–399 (2020).

75. Lubarski-Gotliv, I. *et al.* FXYD5 (Dysadherin) may mediate metastatic progression through regulation of the beta-Na(+)-K(+)-ATPase subunit in the 4T1 mouse breast cancer model. *Am. J. Physiol. Cell Physiol.* **313**, C108–C117 (2017).

76. Gharanei, S. *et al.* RASSF2 methylation is a strong prognostic marker in younger age patients with Ewing sarcoma. *Epigenetics-US.* **8**, 893–898 (2013).

77. Chen, S. *et al.* Prognostic value of GIMAP4 and its role in promoting immune cell infiltration into tumor microenvironment of lung adenocarcinoma. *Biomed. Res. Int.* **2022**, 7440189 (2022).

78. Guo, Z. *et al.* Water channel protein AQP1 in cytoplasm is a critical factor in breast cancer local invasion. *J. Exp. Clin. Cancer Res.* **42**, 49 (2023).
79. Xu, X. *et al.* ITGA5 promotes tumor angiogenesis in cervical cancer. *Cancer Med.* **12**, 11983–11999 (2023).
80. Lee, J. H. *et al.* Induction of squamous cell carcinoma after map3k8 overexpression in murine salivry gland epithelial cells. *Head Neck.* **41**, 924–929 (2019).
81. Liu, Y. & Yin, S. A novel prognostic index based on the analysis of glycolysis-related genes in head and neck squamous cell carcinomas. *J Oncol.* **2020**, 7353874 (2020).
82. Jin, Y. & Qin, X. Significance of TP53 mutation in treatment and prognosis in head and neck squamous cell carcinoma. *Biomark. Med.* **15**, 15–28 (2021).
83. Shi, C., Liu, S., Tian, X., Wang, X. & Gao, P. A TP53 mutation model for the prediction of prognosis and therapeutic responses in head and neck squamous cell carcinoma. *BMC Cancer.* **21**, 1035 (2021).
84. Mori, T. Involvement of the p53–p16/RB pathway control mechanism in early-stage carcinogenesis in head and neck squamous cell carcinoma. *Pathol. Int.* **72**, 577–588 (2022).
85. Wang, L. *et al.* Overexpression of TP53 mutation-associated microRNA-182 promotes tumor cell proliferation and migration in head and neck squamous cell carcinoma. *Arch. Oral Biol.* **73**, 105–112 (2017).
86. Lim, S. M. *et al.* Investigating the feasibility of targeted next-generation sequencing to guide the treatment of head and neck squamous cell Carcinoma. *Cancer Res. Treat.* **51**, 300–312 (2019).

Acknowledgements

Author contributions

Q.L. conceived and designed this study. L.L. and Q.L. collected data and analyzed the data. L.L. wrote the manuscript. Q.L. were responsible for the final modified version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-60516-6.

Correspondence and requests for materials should be addressed to Q.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.