



An integrated bioinformatics analysis of the S100 in head and neck squamous cell carcinoma

Guo Bai^{1,2,3#}, Shijing Yue^{1,2,3#}, Yuanhe You^{2,3,4}, Jiong Zhao^{1,2,3}, Zhuowei Tian^{2,3,4}, Yanan Wang^{2,3,4}, Zhiyuan Zhang^{1,2,3,4}, Chi Yang^{1,2,3^}

¹Department of Oral Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; ²College of Stomatology, Shanghai Jiao Tong University, Shanghai, China; ³National Center for Stomatology, National Clinical Research Center for Oral Diseases, Shanghai Key Laboratory of Stomatology, Shanghai, China; ⁴Department of Oral and Maxillofacial-Head and Neck Oncology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Contributions: (I) Conception and design: C Yang, Z Zhang; (II) Administrative support: C Yang, Z Zhang; (III) Provision of study materials or patients: C Yang, Y Wang; (IV) Collection and assembly of data: G Bai, Y You, Z Tian; (V) Data analysis and interpretation: G Bai, S Yue, J Zhao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Chi Yang. Department of Oral Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, No. 639, Zhizaoju Road, Huangpu District, Shanghai 200011, China. Email: yangchi633@163.com; Zhiyuan Zhang. Department of Oral and Maxillofacial-Head and Neck Oncology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, No. 639, Zhizaoju Road, Huangpu District, Shanghai 200011, China. Email: zhzy0502@163.com.

Background: This study aims to evaluate the expression profile and clinical value of the S100 family in head and neck squamous cell carcinoma (HNSCC).

Methods: The expression patterns, clinicopathological features, prognostic significance and underlying correlations of S100 family genes in HNSCC were determined by bioinformatics analysis with the application of several databases, including the The Cancer Genome Atlas (TCGA) and OncoPrint for differential expression gene (DEG) analysis, and a series of analysis tools, including Database for Annotation, Visualization and Integrated Discovery (DAVID), cBioPortal, Kaplan-Meier Plotter, Tumor Immune Estimation Resource (TIMER) and R software packages.

Results: The results from the study demonstrated that S100A4, S100A10, and S100A13 may act as prognostic markers through overall survival (OS), disease-free survival (DFS) and tumor infiltrating immune cell enrichment and a prognostic S100 family gene model comprising *S100A1–A4*, *S100A8*, *S100A10*, *S100A12*, and *S100A13* was identified. The messenger RNA (mRNA) expression of S100A1, S100A9, S100A14, and S100A7A was significantly different in HNSCC patients, together with a high mutation rate of the S100 family was found. Evaluation of clinicopathological value demonstrated the heterogeneity of S100 family functions. S100A1, S100A7, S100A8, S100A9, S100A13, S100A14, and S100A16 were observed to significantly correlate with multiple biological processes (BPs) of HNSCC, including initiation, lymph node metastasis, and lymphovascular invasion. In addition, the S100 family were significantly associated with epithelial-mesenchymal transition (EMT)-related genes.

Conclusions: This present study demonstrated that S100 family members are implicated in the initiation, progression, metastasis and survival of HNSCC.

Keywords: Head and neck squamous cell carcinoma (HNSCC); S100 family; bioinformatics analysis

Submitted May 13, 2022. Accepted for publication Mar 07, 2023. Published online Apr 12, 2023.

doi: 10.21037/tcr-22-1353

View this article at: <https://dx.doi.org/10.21037/tcr-22-1353>

[^] ORCID: 0000-0001-9926-7431.

Introduction

Over 600,000 people worldwide are annually diagnosed with head and neck cancer (1,2). Head and neck squamous cell carcinoma (HNSCC) encompasses a group of biologically similar cancers arising from multiple anatomic subsites in the head and neck region, including the oral cavity, lip, larynx and hypopharynx (3). Despite recent advancements in treatment modalities for HNSCC, the long-term survival rates (40–50% survival for <5 years) have not been significantly improved over the past decade (4,5). Given the high mortality of HNSCC, it is vital to identify early diagnostic and prognostic biomarkers.

The S100 family consists of a group of low molecular weight (10,000 Da) proteins with two calcium-binding sites that have helix-loop-helix EF-hand type conformations (6). At least 25 closely related members have been identified in humans (7). Although S100 family members exhibit a high structural similarity, they are not functionally interchangeable and are implicated in many intracellular biological processes (BPs) such as proliferation, apoptosis, migration, differentiation and cell growth (6,8). In contrast to their intracellular functions, extracellular S100 proteins act in a cytokine-like manner by binding to cell surface receptors, such as the receptor for advanced glycation endproducts (RAGEs) and Toll-like receptors (TLRs) (9). Also, S100s have been revealed to promote tumor

progression by regulating tumor infiltrating immune cells and tumor immune response (10,11). Furthermore, previous studies have reported that S100 proteins are associated with the regulation of epithelial-mesenchymal transition (EMT), cancer stem cells (CSCs) and tumor heterogeneity in human carcinoma (12–15).

Due to the wide range of functions of S100 proteins, quantification and recognition of S100 family genes may contribute to the consideration of S100 proteins as biomarkers for the early diagnostic and prognosis, as well as targets for therapy. However, there is a lack of reports in analyzing S100 family members in HNSCC based on large cohort data, and the clinical value of these molecules in HNSCC remains to be ascertained. Therefore, the present study aimed to evaluate the expression profile and clinical value of the S100 family in HNSCC using a series of online databases. We present the following article in accordance with the STREGA reporting checklist (available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-1353/rc>).

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Oncomine analysis

Oncomine is an online cancer gene expression microarray database that can be used to analyze S100 transcription levels in different cancers (<http://www.oncomine.org/resource/login.html>) (16). In this study, S100 messenger RNA (mRNA) expression in clinical cancer tissues with that in normal control tissues was compared. P value and fold change cutoffs were 0.05 and 2, respectively.

Differential expression gene analysis

Transcriptome RNA-sequencing (RNA-seq) data of HNSCC samples was obtained from the Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>) (17). Raw count data and relevant clinical information for the patients were also downloaded and extracted. The method of acquisition and application complied with the guidelines and policies. The differential expressions of S100 family proteins between the tumor tissues and normal tissues were analyzed using the Wilcoxon test and visualized by the R software (v4.0.3) and R package ggplot2 (v3.3.2). A P value

Highlight box

Key findings

- S100 family proteins are significantly associated with the development of HNSCC.

What is known and what is new?

- S100s could promote tumor progression by regulating tumor infiltrating immune cells and tumor immune response. They are associated with regulation of EMT, CSCs and tumor heterogeneity in human carcinomas.
- The mRNA expression of S100A1, S100A9, S100A14, and S100A7A was significantly different in HNSCC patients compared with other cancers. A prognostic S100 family gene model comprising S100A1–A4, S100A8, S100A10, S100A12, and S100A13 was identified in HNSCC. S100A4, S100A10, and S100A13 may act as prognostic markers.

What is the implication, and what should change now?

- A prognostic signature of the eight S100 genes could moderately predicted the prognosis of HNSCC patients. Further investigations are necessary to explore and verify the underlying molecular mechanisms.

of less than 0.05 was considered statistically significant.

Genes expression association

The dataset used comprised HNSCC mRNA-seq data from TCGA. A heat map of the correlation between S100 family genes and EMT-related genes of *SNAIL1* (18), *VIM* (19), *TWIST* (19), *CDH1* (20), *CDH2* (20) were performed by the Spearman correlation via R software package ggstatsplot (21), and is displayed by the R software package pheatmap (21). A P value of less than 0.05 was considered statistically significant.

Enrichment analysis of S100s family in HNSCC

The function enrichment analysis for gene ontology (GO) terms, including the BP, cellular component (CC), and molecular function (MF) categories were performed in the Database for Annotation, Visualization and Integrated Discovery (DAVID; version 6.8, <https://david.ncifcrf.gov/tools.jsp>) (22). The analysis threshold was determined by the false discovery rate (FDR) value, and FDR <0.05 for BP, FDR <0.17 for CC, and FDR <0.01 for MF indicate that the functional comment is significantly enriched.

Analysis of gene alteration frequency

cBioPortal (<https://www.cbioportal.org/>) is an open-access online resource for exploring and visualizing multidimensional cancer genomics dataset (23). HNSCC (TCGA, provisional) dataset was selected to analyze genetic alterations of S100 family genes in HNSCC. The genomic profiles included mutations, putative copy-number alterations, and mRNA expressions (RNA-seq V2 RSEM with z-scores =±2).

Kaplan-Meier plotter

The Kaplan-Meier plotter (<https://www.kmplot.com/>) is an online database comprises gene expression information and clinical outcome parameters of liver cancer, breast cancer, lung cancer and various cancer types (24). To analyze correlation of S100 family members with overall survival (OS) and disease-free survival (DFS) rates of HNSCC patients, samples were divided into two groups (low and high expression) based on median mRNA levels with a hazard ratio (HR) with 95% confidence intervals (CIs) and log-rank P values. Log-rank P values <0.05 were considered

as significant.

Establishment of prognostic S100s family gene model

Raw counts of RNA-seq data (level 3) and corresponding clinical information from 501 HNSCC patients were obtained from TCGA (<https://portal.gdc.cancer.gov/>). The least absolute shrinkage and selection operator (LASSO) regression algorithm for feature selection, using 10-fold cross-validation, the above analysis uses the R software package glmnet (v4.1-1).

Multifactor Cox regression was applied to analyze the 19 S100 proteins, and a risk model of 8 S100s family genes was ultimately established. The individualized risk score with coefficients was calculated to construct a prognostic signature separating the high-risk and low-risk groups. According to the median risk score, subgroups including low- and high-risk groups were established (25). For Kaplan-Meier curves, P values and HR with 95% CI were generated by log-rank tests and univariate Cox proportional hazards regression. timeROC (v0.4) analysis was performed to compare the predictive accuracy of each gene and risk score. All analytical methods above and R packages were performed using R software version 4.0.3 (The R Foundation for Statistical Computing, 2020). P<0.05 was considered as statistically significant.

Tumor-infiltrating immune cell scores based on the risk model

The correlation between signature risk score and the tumor-infiltrating immune cell (including B cells, CD4⁺ T cells, CD8⁺ T cells, macrophages, neutrophils, and dendritic cells) was performed using the Tumor Immune Estimation Resource (TIMER) database, a network resource that systematically evaluates the clinical effects of different immune cells in different types of cancer (<http://cistrome.dfci.harvard.edu/TIMER/>) (26). In the above analyses, P<0.05 was considered statistically significant.

Clinicopathological features evaluation

Genomic and clinical annotation data based on the TCGA primary HNSCC cohort were filtered for clinicopathological analysis. The ratio of neoplasm histologic grade (G1 + G2/G3 + G4), tumor size (T1 + T2/T3 + T4), nodal status (N0/N0+), perineural invasion (no/yes), pathological nodal extracapsular spread (no/yes),

lympho-vascular invasion (no/yes), the relationship between new tumor event (no/yes) after initial treatment and S100s expression (low/high) were evaluated.

Statistical analysis

Statistical analyses were carried out using SPSS Statistics 20.0 software (IBM SPSS, Chicago, IL, USA). The statistical differences between the experimental groups were analyzed by Student's *t*-test. The relationship between S100s expression and survival rates was analyzed using Kaplan-Meier curve. Clinicopathological features were evaluated by crosstab analyses and chi-squared tests were performed. Statistically, $P < 0.05$ and log-rank P values < 0.05 were considered statistically significant.

Results

Atlas of selected 19 S100 protein members in HNSCC

There are 21 S100 proteins, which are exclusively found in vertebrates, encoded in the human genome. The genes encoding 17 of 21 family members (S100A1–A14, S100A7A, S100A7L2 and S100A16) are located on chromosome 1q21, which is referred to as the epidermal differentiation complex (EDC) (27). The genes of the remaining 4 family members, namely, *S100B*, *S100G*, *S100P*, and *S100Z* are dispersed throughout the genome on chromosomes 21, X, 4, and 5, respectively. *S100G* and *S100A7L2* were excluded from this study because their detailed data were not sufficient for a comprehensive analysis. The full names and biological functions of the 19 included family members (S100A1–A14, S100A7A, S100A16, S100P, S100B, and S100Z) are summarized in Table S1.

Transcriptional expression of the S100 protein family in HNSCC patients

First, the transcription expression changes of individual S100 proteins in HNSCC compared to normal tissues were analyzed using the Oncomine database (Figure 1A). S100A2, S100A3, S100A7, S100A7A, and S100A10 were mainly upregulated in HNSCC, while S100A1, S100A4, S100A6, S100A8, S100A9, S100A10, S100A12, S100A14, and S100P were downregulated in HNSCC. Furthermore, the mRNA expression level of S100 proteins between HNSCC and normal tissues was compared. As shown in Figure 1B, the expression of S100A8, S100A9, and S100A14

was significantly lower in HNSCC tissues compared to normal control tissues, while the expression of S100A2, S100A3, S100A7A, and S100A10 was higher in HNSCC tissues compared to normal control tissues ($P < 0.01$), which was in accordance with S100 expression changes. The genetic alteration of S100 family members in HNSCC was then analyzed (Figure S1A, S1B). As shown in Figure S1A, high mutation rates of the S100 family were observed in HNSCC patients (22.54%). Among all types of mutations, mRNA upregulation was the most common alteration (14.75%). The prognostic roles of gene alterations in patients with HNSCC were also analyzed, however, no significant correlations were observed between the presence of alterations and OS/DFS ($P = 0.135$ and $P = 0.781$, respectively, Figure S1C). In addition, correlations among S100 family members in HNSCC were verified (Figure S2). The expression of S100A7, S100A7A, S100A8, and S100A9 was positively correlated with each other. The expression of S100A16 was positively related with S100A14 and S100A2. Moreover, the expression of S100A4, S100A5, and S100A13 was negatively related to the expression of S100A7, S100A7A, S100A8, and S100A9.

Predicted functions of S100 proteins in HNSCC

According to the GO enrichment analysis, the cell components (CCs) of the S100 family were mainly enriched in the nucleus, extracellular exosome and cytoplasm (Figure 2A). RAGE receptor binding, protein binding and calcium ion binding were the three main enriched MFs of S100 proteins in HNSCC (Figure 2B). The enriched BPs of S100 proteins included the innate immune response, defense response to bacteria and fungi, positive regulation of the inflammatory response and positive regulation of I- κ B kinase/nuclear factor- κ B (NF- κ B) signaling (Figure 2C).

Association between S100 mRNA expression and HNSCC prognosis

Kaplan-Meier plotter curves and log-rank test analyses were performed to explore the effect of S100 mRNA expression on the survival of HNSCC patients. Higher expression of S100A10 and S100A13 indicated significantly poorer OS ($P < 0.05$). In contrast, lower expression of S100A4 was associated with significantly poorer OS ($P < 0.05$) (Figure 3). The average of the values high and low S100s mRNA expression was shown in Table S2. In addition, only higher expression of S100A13 indicated significantly worse DFS for HNSCC patients ($P < 0.001$) (Figure 4).

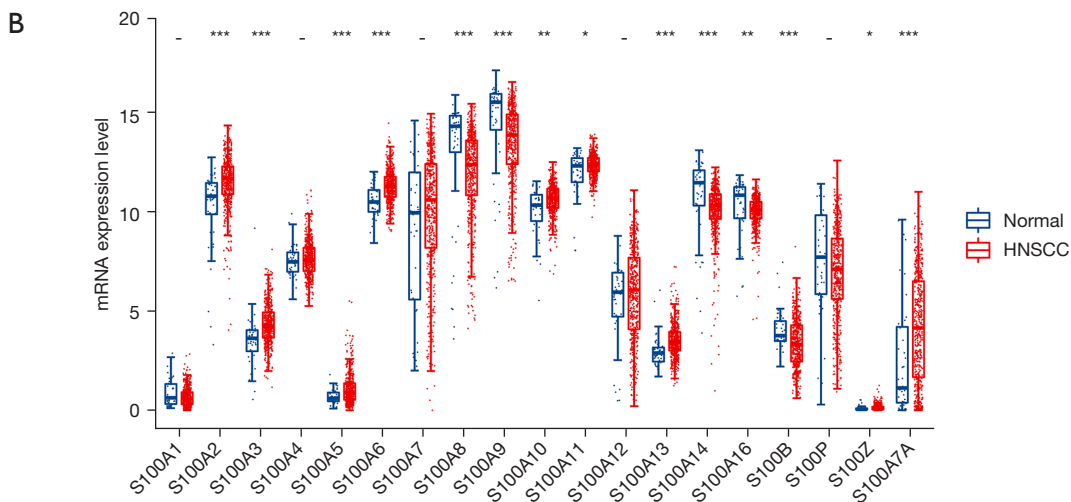
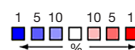
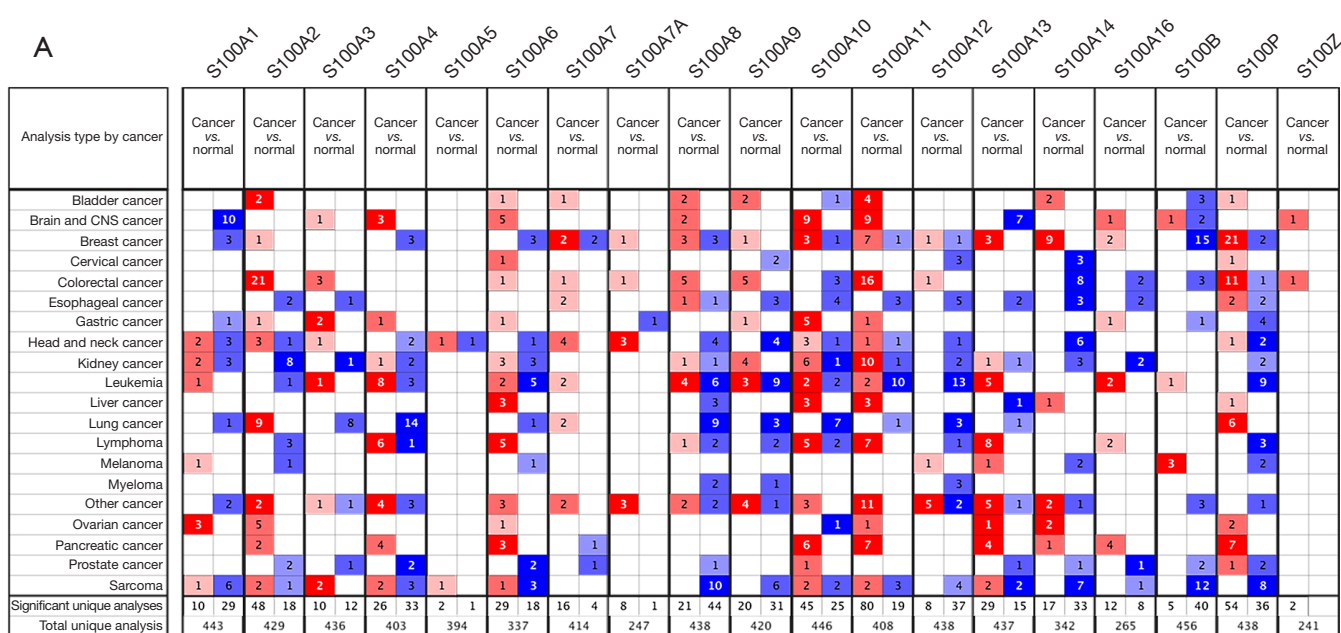


Figure 1 Transcription expression of S100 proteins in various cancers and HNSCC. (A) Number of datasets with statistical significance for S100s mRNA expression in different types of cancers: red, upregulated; blue, downregulated. (B) mRNA expression distribution of S100 proteins in HNSCC tumor tissues and normal tissues. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; -, not significant; CNS, central nervous system; mRNA, messenger RNA; HNSCC, head and neck squamous cell carcinoma.

The clinicopathological value of the S100 protein family based on the TCGA HNSCC cohort is summarized in *Table 1*. S100A2, S100A7, S100A9, S100A12, S100A14, S100A16, and S100A7A were identified as significant for the evaluation of neoplasm histologic grade and nodal invasion. For evaluation of new tumor events after initial treatment,

S100A13 was identified as significant.

An eight-gene prognostic signature of S100s in HNSCC

To further identify the S100 proteins that are significantly correlated with the prognosis of HNSCC patients, LASSO

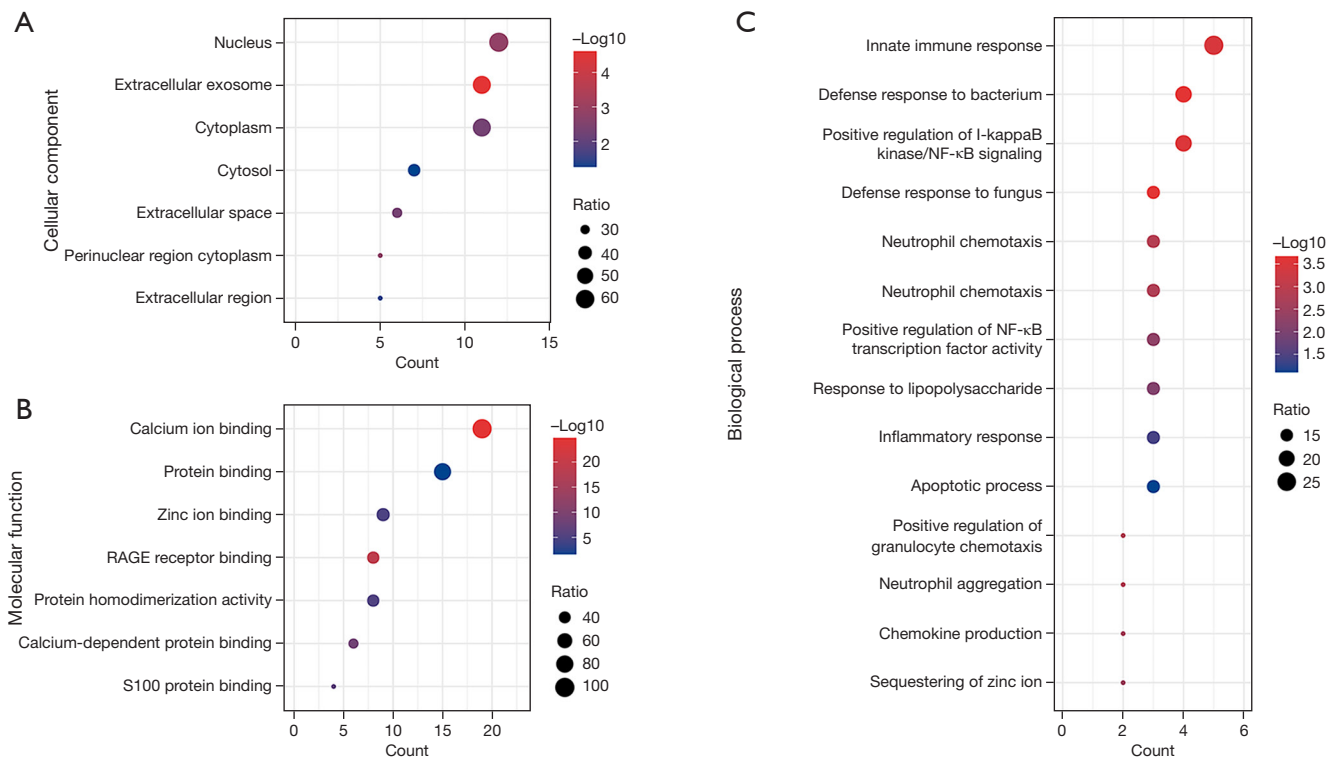


Figure 2 GO enrichment analysis of S100s in HNSCC. (A,B) The top most 7 enriched enrichment pathways in CCs and MFs, respectively. (C) The top most 14 enriched GO enrichment pathways in BPs. RAGE, receptor for advanced glycation endoprotect; NF- κ B, nuclear factor- κ B; GO, gene ontology; HNSCC, head and neck squamous cell carcinoma; CC, cellular component; MF, molecular function; BP, biological process.

regression with tenfold cross-validation was performed. Under the optimal lambda value ($\lambda_{\min} = 0.0148$), 8 genes that were significantly associated with OS were identified among the 19 S100 genes (Figure 5A, 5B). Thereafter, the eight gene-based risk score was calculated based on their Cox coefficients as follows: risk score = $-0.1053 \times$ (expression level of *S100A1*) + $-0.1339 \times$ (expression level of *S100A2*) + $-0.1516 \times$ (expression level of *S100A3*) + $-0.2457 \times$ (expression level of *S100A4*) + $-0.0263 \times$ (expression level of *S100A8*) + $0.4372 \times$ (expression level of *S100A10*) + $-0.0121 \times$ (expression level of *S100A12*) + $0.1782 \times$ (expression level of *S100A13*). Based on the risk scores, the patients were divided into the high-risk group (n=251) and low-risk group (n=252) (Figure 5C). Figure 5D shows the survival status of all patients in the training group and Figure 5E presents the heatmap of the eight prognostic genes. The Kaplan-Meier survival curves showed that the high-risk group had worse OS than the low-risk group (Figure 5F). Moreover, the eight-gene prognostic signature showed moderate prediction ability in a time-dependent receiver operating

characteristic (ROC) analysis (Figure 5G) for 1-, 3-, and 5-year OS.

Negative correlation of S100 family members with tumor-infiltrating immune cells in HNSCC patients

The association of the mRNA expression status of the eight-gene prognostic signature with tumor-infiltrating immune cells was identified by TIMER in HNSCC patients. As illustrated in Figure 6, a negative correlation was found between S100 expression and all six types of immune cells (B cells, CD4⁺ T cells, CD8⁺ T cells, macrophages, neutrophils, and dendritic cells), thereby revealing prognostic value between S100 family mRNA expression and immune cell infiltration in HNSCC patients.

Correlation between S100 family proteins and EMT- and CSC-related genes, PD-1 and PDL1 genes

EMT is a vital bioprocess for the invasion and metastasis

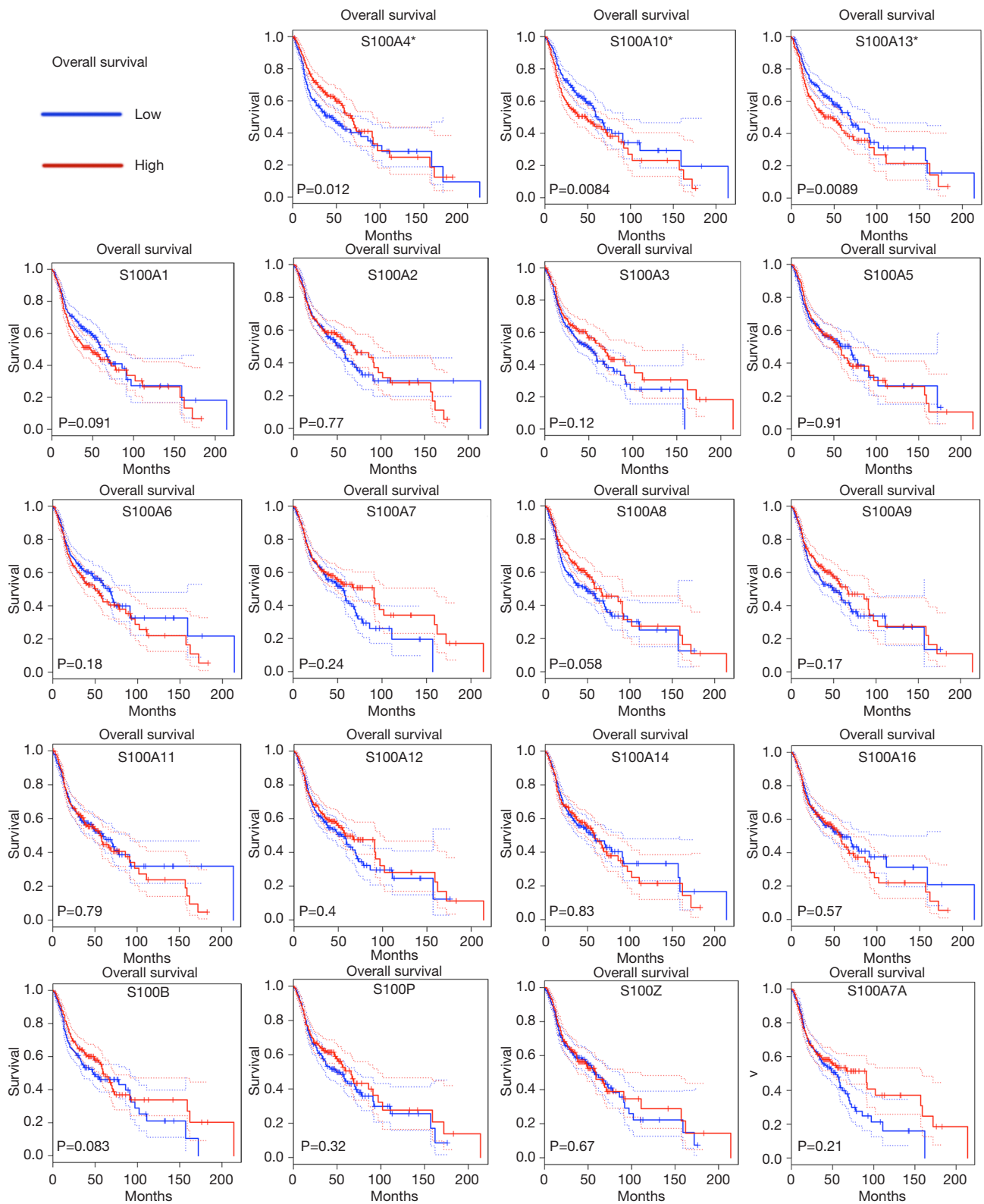


Figure 3 Prognostic value of S100s for the OS of HNSCC patients. The expression of S100A4, S100A10, and S100A13 were significantly associated with OS. *, $P < 0.05$. OS, overall survival; HNSCC, head and neck squamous cell carcinoma.

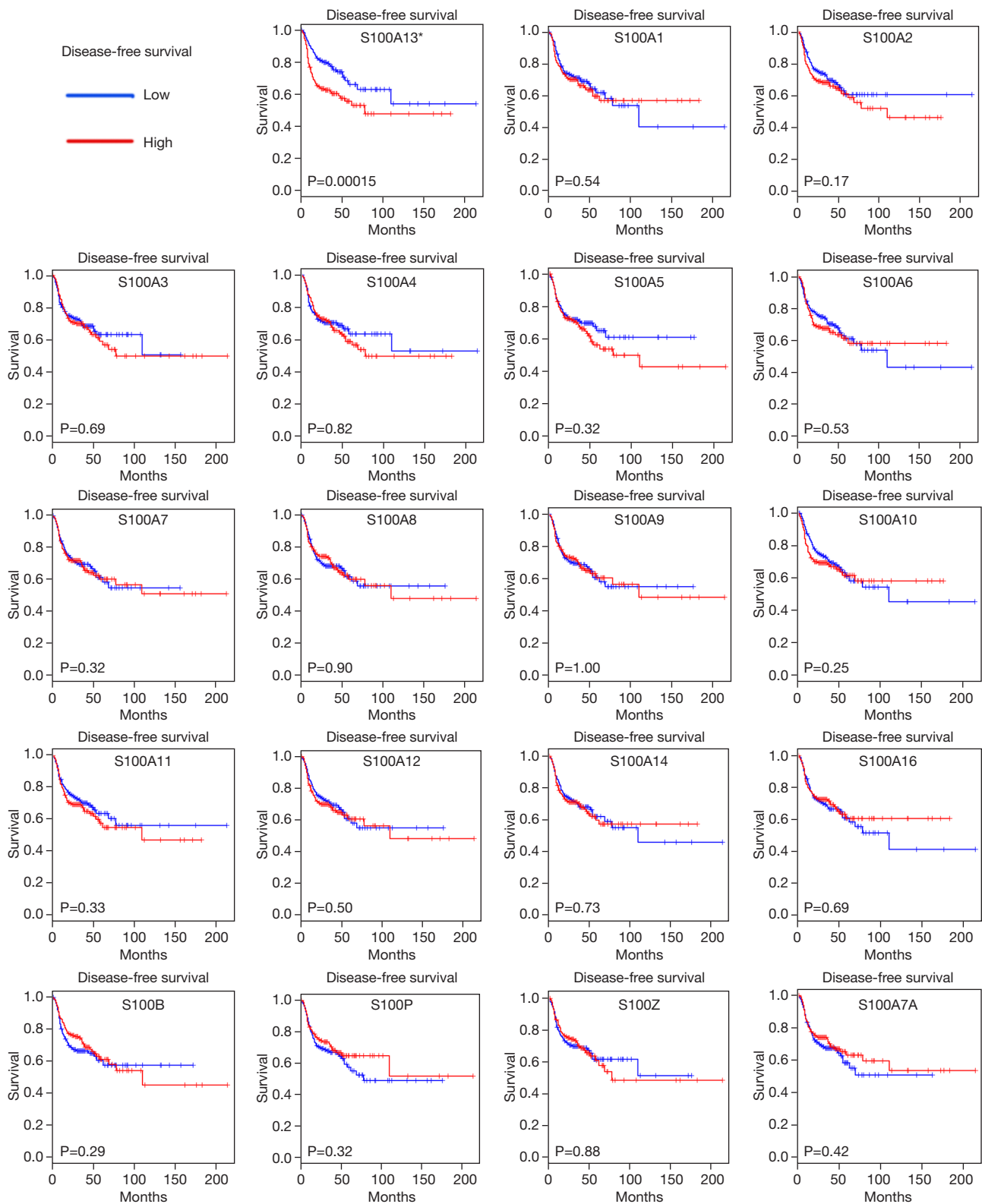


Figure 4 Prognostic value of S100s for the DFS of HNSCC patients. The expression of S100A13 was significantly associated with DFS. *, $P < 0.001$. DFS, disease-free survival; HNSCC, head and neck squamous cell carcinoma.

Table 1 Clinicopathological evaluation for the reported CSC biomarkers based on TCGA primary HNSCC cohort

S100 protein family	Neoplasm histologic grade (G1 + G2/G3 + G4), P value	Tumor size (T1 + T2/T3 + T4), P value	Nodal status (N0/N0+), P value	Perineural invasion (no/yes), P value	Pathological nodal extracapsular spread (no/yes), P value	Lympho-vascular invasion (no/yes), P value	New tumor event after initial treatment (no/yes), P value
S100A1	0.676	0.040 ^a	0.395	0.006 ^a	0.053	0.018 ^a	0.758
S100A2	<0.001 ^{ab}	0.117	0.017 ^a	0.419	0.061	0.020 ^{ab}	0.303
S100A3	0.168	0.128	0.167	0.270	0.308	0.264	0.570
S100A4	0.364	0.093	0.410	0.436	0.869	0.838	0.972
S100A5	0.171	0.267	0.436	0.336	0.896	0.282	0.733
S100A6	0.199	0.284	0.341	0.628	0.963	0.332	0.506
S100A7	<0.001 ^{ab}	0.540	0.013 ^a	0.244	0.013 ^{ab}	<0.001 ^{ab}	0.843
S100A8	<0.001 ^{ab}	0.147	0.662	0.905	0.254	<0.001 ^{ab}	0.808
S100A9	<0.001 ^{ab}	0.208	0.001 ^{ab}	0.289	0.070	<0.001 ^{ab}	0.401
S100A10	0.066	0.066	0.110	0.089	0.422	0.901	0.198
S100A11	0.001 ^{ab}	0.938	0.632	0.478	0.779	0.747	0.168
S100A12	<0.001 ^{ab}	0.093	0.001 ^{ab}	0.709	0.001 ^{ab}	<0.001 ^{ab}	0.982
S100A13	0.285	0.910	0.058	0.135	0.020	0.149	0.013 ^a
S100A14	<0.001 ^{ab}	0.157	0.461	0.289	0.004 ^{ab}	0.026 ^{ab}	0.207
S100A16	<0.001 ^{ab}	0.378	0.011 ^{ab}	0.234	0.009 ^{ab}	0.01 ^{ab}	0.093
S100B	0.047 ^a	<0.001 ^{ab}	0.613	0.401	0.477	0.582	0.874
S100P	0.116	0.999	0.126	0.532	0.105	0.464	0.544
S100Z	0.304	0.118	0.532	0.950	0.943	0.695	0.544
S100A7A	<0.001 ^{ab}	0.569	0.001 ^{ab}	0.860	0.001 ^{ab}	<0.001 ^{ab}	0.295

*, P<0.05; ^a, positive correlation with significance; ^b, negative correlation with significance. CSC, cancer stem cell; TCGA, The Cancer Genome Atlas; HNSCC, head and neck squamous cell carcinoma.

of HNSCC. Information on the underlying correlation between S100 family proteins and EMT-related genes is rare. The heatmaps shown in [Figure S3](#) elaborate the expression patterns of EMT-related genes (*SNAIL*, *VIM*, *TWIST1*, *CDH2*, and *CDH1*). Most of the S100 family members showed a significantly negative correlation with EMT-related genes, such as S100A8, S100A9, and S100A14. The same as in correlation with CSC-related genes ([Figure S4](#)). Programmed death 1 (PD-1; CD279) exerts an important role in immune homeostasis (28) and the expression of programmed death ligand 1 (PD-L1) has been identified in HNSCC (29). The relative mRNA expression levels of both S100s and PD-1, PD-L1 in HNSCC were analyzed ([Figure S5A,S5B](#)). [Table 2](#) shows the expression relationship between S100s members and related genes, and

P<0.05 was considered as significant correlation. Only PD-L1 showed significantly associated with OS rate in HNSCC patient (P=0.04937), as is shown in [Figure S5C](#).

Discussion

Dysregulated S100 expression is a common feature in several human cancers, and each cancer exhibits a distinctive S100 protein profile. Emerging evidence indicates that the biology of most S100 proteins is complex and multifactorial, and these proteins actively contribute to tumorigenesis, cancer progression and metastasis (6,8). Great progress has been made since the past few years, but systematic analysis of the clinical value of S100 proteins in HNSCC is still inadequate. Here, the roles of 19 S100 protein families

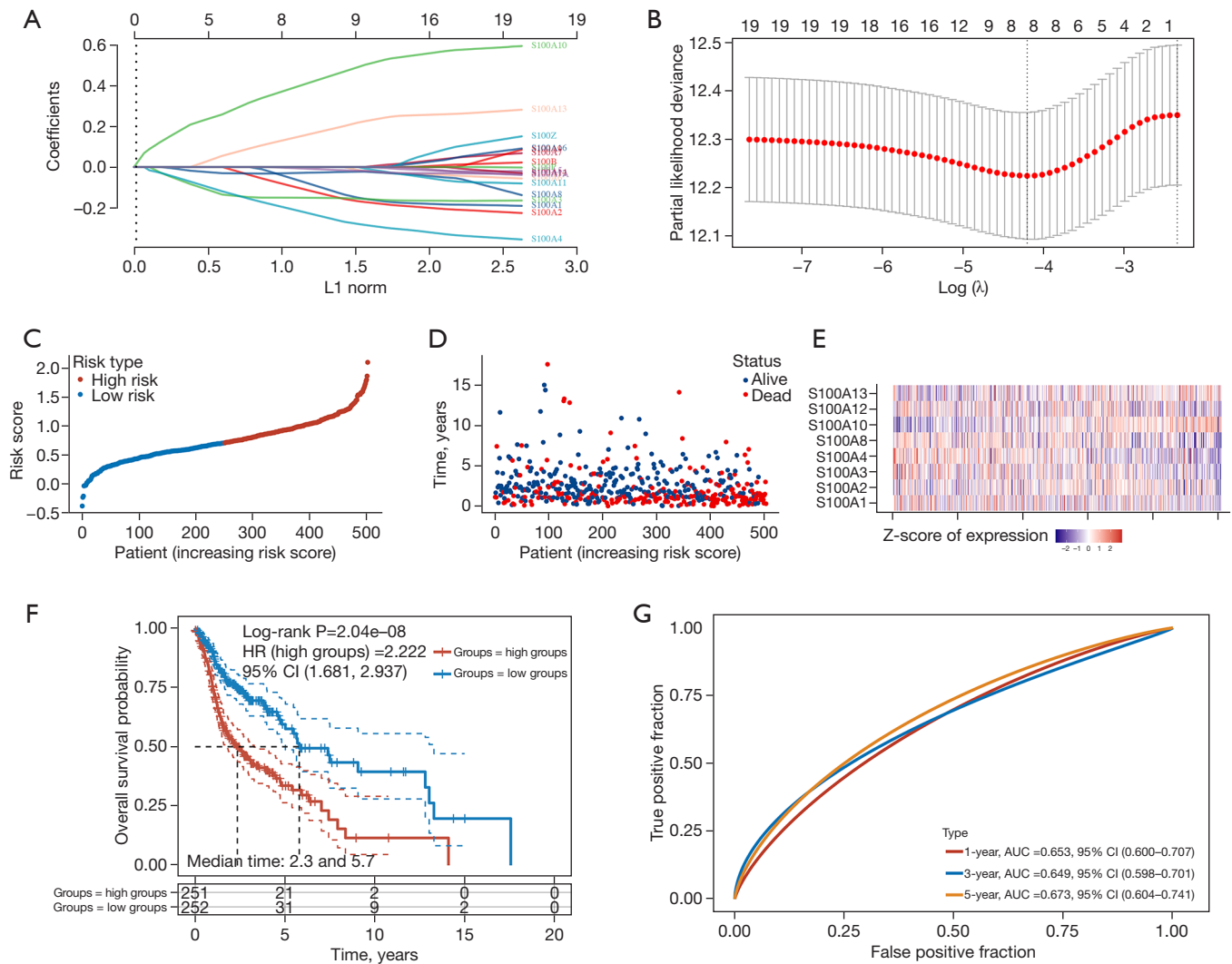


Figure 5 Identification of 8 significantly prognostic S100 genes and prognostic analysis of eight-gene signature. (A) The LASSO coefficient profile of 19 S100 family protein-coding genes. (B) LASSO regression with tenfold cross-validation obtained 8 prognostic genes using minimum lambda value. (C) The dotted curve of risk score and patients were divided into low- and high-risk group. (D) Survival status of the patients. More dead patients corresponding to the higher risk score. (E) Heatmap of the expression profiles of the eight prognostic genes in low- and high-risk group. (F) Kaplan-Meier survival analysis of the eight-gene signature. (G) Time-dependent ROC analysis the of the eight-gene signature. HR, hazard ratio; CI, confidence interval; AUC, area under the curve; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic.

were comprehensively evaluated through bioinformatics analysis based on over 500 primary HNSCC cases from the TCGA cancer database. Differential expression gene (DEG) analysis showed differences in the expression of S100 family members in HNSCC, indicating that S100 proteins play different but important roles in HNSCC development. The functional enrichment analysis results suggested that

S100 proteins might play an important role in the immune response during HNSCC via protein binding and NF- κ B related signaling pathways.

The clinicopathological features of HNSCC patients suggested that S100 proteins may regulate the occurrence and progression processes through different mechanisms, warranting further attention to demonstrate the underlying

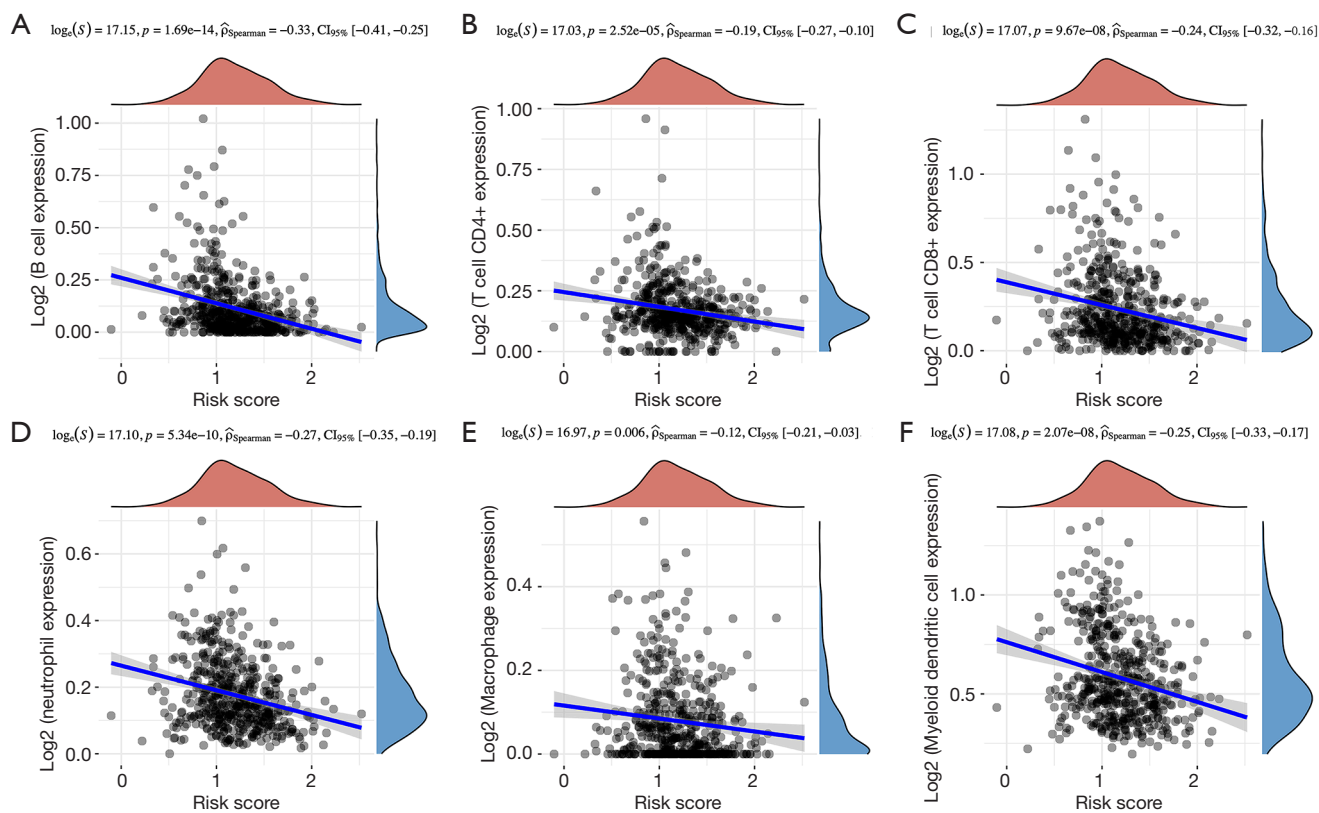


Figure 6 Negative relationships between the eight-gene prognostic signature and infiltration abundances of six types of immune cells. (A) B cells. (B) CD4 T cells. (C) CD8 T cells. (D) Neutrophils. (E) Macrophages. (F) Dendritic cells. CI, confidence interval.

heterogeneity. Eight *S100* genes, namely, *S100A1–A4*, *S100A8*, *S100A10*, *S100A12*, and *S100A13* were identified for their significant association with HNSCC patient prognosis. The combined effect of these *S100*s were calculated. As an eight-gene prognostic signature, *S100* proteins had a positive correlation with HNSCC patient survival with moderate prediction capability. Additionally, the *S100* family-based signature manifested was closely related to the immune microenvironment, showing a negative relationship with different types of tumor-infiltrating immune cells.

The expression profile and functions of *S100* family proteins are quite different in HNSCC from that in other tumors. *S100A7* has been found in various epithelial malignancies, including HNSCCs, and is associated with cancer initiation and poor prognosis. Inflammatory cytokines regulate *S100A7* expression in cancers (30), implicating their roles in the tumor microenvironment. The *S100A7* gene, together with NF- κ B transcription factors, plays an important role in inducing cell growth

by stimulating proliferation genes (31). As an oncogene, *S100A7* is the major contributing factor in the occurrence of oral cancer and promotes local tumor progression by activating the mitogen-activated protein kinase (MAPK) signaling pathway via the Ras-related protein 2A (RAB2A) pathway (12). In addition, *S100A7* promotes angiogenesis by regulating matrix metalloproteinase (MMP)13 and vascular endothelial growth factor (VEGF) (32). However, in the present study, *S100A7* transcript expression in HNSCC showed no significant difference compared to normal tissue. Thus, *S100A7* was not included in the prognostic signature. *S100A8* and *S100A9* are often co-expressed and form a heterodimeric complex named as calprotectin. *S100A8/A9* proteins are important constituents of the tumor microenvironment that critically contribute to the progression of tumors (33). Moreover, *S100A8/A9* proteins work as inflammatory chemoattractant to recruit inflammatory cells to damaged tissues, thereby contributing to tumorigenesis and cancer metastasis (34). However, the results in the present study showed that *S100A8/A9* were

Table 2 Correlation of S100 family members with PD-1 and PD-L1 of HNSCC patients

S100 family	PDCD1, P value	CD274, P value
S100A1	0.079	0.219
S100A2	<0.001 ^{ab}	0.792
S100A3	0.861	0.598
S100A4	<0.001 ^{aa}	0.725
S100A5	0.114	0.002 ^{ab}
S100A6	0.035 ^{ab}	0.094
S100A7	0.014 ^{ab}	0.187
S100A8	0.860	0.064
S100A9	0.429	0.017 ^{aa}
S100A10	<0.001 ^{ab}	0.253
S100A11	0.014 ^{ab}	0.792
S100A12	0.054	0.930
S100A13	0.219	0.018 ^{ab}
S100A14	0.009 ^{ab}	0.014 ^{ab}
S100A16	0.011 ^{ab}	0.077
S100B	<0.001 ^{aa}	0.002 ^{aa}
S100P	0.565	0.334
S100Z	<0.001 ^{aa}	0.002 ^{aa}
S100A7A	0.114	0.860

^{*}, P<0.05; ^a, positive correlation with significance; ^b, negative correlation with significance. PD-1, programmed death 1; PD-L1, programmed death ligand 1; HNSCC, head and neck squamous cell carcinoma.

downregulated in HNSCC with a negative correlation with the expression of EMT genes, supporting the notion that S100A8/A9 play multifaceted tumor-suppressive roles in HNSCC (14). Decreased expression of S100A8 and S100A9 has been reported to be associated with poor prognosis in oral carcinoma (35). At low concentrations, S100A8/A9 promotes tumor cell growth by activating the MAPK and NF- κ B signaling pathways, depending on RAGE ligation (36). Silencing endogenous S100A8 and S100A9 expression in TR146 cells increases MMP2 activity, thereby promoting cell invasion and migration (37). Cytoplasmic calprotectin also functions to control the cell cycle DNA damage checkpoint at G₂/M and block proliferation (38). S100A14 expression is associated with cancer differentiation, lymph node status and metastasis (39). High expression

of S100A14 acts as a mediator of EMT and regulates the proliferation, migration and invasion of human cervical and breast cancer cells (40,41). In contrast, S100A14 suppresses nasopharyngeal carcinoma (NPC) metastasis by inhibiting the NF- κ B signaling pathway and reversing EMT. Lower S100A14 expression is significantly correlated with shorter patient OS and distant metastasis-free survival in NPC patients (42), same in oral squamous cell carcinoma (OSCC) (43-45). Nevertheless, the analysis in the present study also supported that the expression of S100A14 is decreased in HNSCC and that low expression of this protein is associated with pathological nodal extracapsular spread and lymphovascular invasion.

S100A4, S100A10, and S100A13 are associated with the OS of HNSCC, and S100A13 is correlated with DFS, according to our analysis. S100A4, also known as metastasin, plays an important role in the invasion and metastasis of human malignant tumors and is considered an EMT marker for being a direct target of β -catenin/T-cell factor (TCF) (46-49). S100A10 upregulation has been reported in a number of cancers and is generally associated with a poor prognosis (50,51). S100A13 plays a key role in tumor growth and metastasis. High expression of S100A13 has been reported to be involved in tumor growth, metastasis and poor OS in thyroid tumors, melanoma and early-stage non-small cell carcinoma (52-54). However, the results in the present study indicated that HNSCC patients with high expression of S100A13 have a better prognosis. At the same time, S100A13 was found to be significantly associated with new tumor events after initial treatment, indicating that S100A13 is a potential biomarker for HNSCC recurrence. The other S100 proteins also showed clinical and pathological significance in HNSCC. Overexpression of S100A1 enhances the cancer cell proliferation, migration and lymph node metastasis in cervical cancer (28), breast cancer (29), and hepatocellular carcinoma (31). In contrast, the expression level of S100A1 in HNSCC tissues did not show any significantly change compared with that in normal tissues. S100A2 expression is downregulated in many cancers, and loss of S100A2 expression is associated with poor prognosis (31). Hence, S100A2 may be a tumor-suppressing protein that potentiates p53 by binding to the p53 transactivation domain, presumably contributing to restoring p53 function in cell arrest and apoptosis (55-59). However, S100A2 is upregulated in some cancers and acts as a tumor promotor (35). High S100A16 expression has been found to be significantly associated with poor prognosis in lung cancer (60) and prostate cancer (61). On the contrary,

S100A16 has been shown to be downregulated in HNSCC and to be associated with malignant phenotypes in OSCC. High S100B expression predicts a good OS in patients with ER-negative breast cancer and good distant metastasis-free survival in patients with breast cancer (62). The results in the present study confirmed that S100B is related to lymph node metastasis and has a significant effect on tumor size.

Conclusions

The results herein indicated that high mutation rate of the S100 family was observed in HNSCC. The mRNA expression of S100A1, S100A9, S100A14, and S100A7A was significantly different in HNSCC patients. In addition, the S100 family was significantly associated with EMT-related genes. For clinical value evaluation, multiple functions of the S100 family were demonstrated. For example, S100A1, S100A7, S100A8, S100A9, S100A13, S100A14, and S100A16 were observed to significantly correlate with initiation, lymph node metastasis, and lymphovascular invasion of HNSCC. In the present study, S100A4, S100A10, and S100A13 may act as prognostic markers. A prognostic signature of eight S100 genes, including *S100A1–A4*, *S100A8*, *S100A10*, *S100A12*, and *S100A13*, moderately predicted the prognosis of HNSCC patients. Considering the heterogeneity of S100 family members, further investigations are necessary to explore and verify the underlying molecular mechanisms.

Acknowledgments

Funding: This work was supported by the National Natural Science Foundation of China (Nos. 81870785 and 81801039), the Science and Technology Commission of Shanghai Municipality Science Research Project (No. 18DZ2291100), and the CAMS Innovation Fund for Medical Sciences (CIFMS) (Project No. 2019-I2M-5-037).

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-1353/rc>

Peer Review File: Available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-1353/prf>

Conflicts of Interest: All authors have completed the ICMJE

uniform disclosure form (available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-1353/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359-86.
2. Siegel RL, Miller KD, Fuchs HE, et al. Cancer Statistics, 2021. *CA Cancer J Clin* 2021;71:7-33.
3. Chow LQM. Head and Neck Cancer. *N Engl J Med* 2020;382:60-72.
4. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer* 2018;18:269-82.
5. Ringash J. Survivorship and Quality of Life in Head and Neck Cancer. *J Clin Oncol* 2015;33:3322-7.
6. Donato R, Cannon BR, Sorci G, et al. Functions of S100 proteins. *Curr Mol Med* 2013;13:24-57.
7. Bresnick AR, Weber DJ, Zimmer DB. S100 proteins in cancer. *Nat Rev Cancer* 2015;15:96-109.
8. Chen H, Xu C, Jin Q, et al. S100 protein family in human cancer. *Am J Cancer Res* 2014;4:89-115.
9. Donato R. Intracellular and extracellular roles of S100 proteins. *Microsc Res Tech* 2003;60:540-51.
10. Wang LJ, Lv P, Lou Y. Alarm Signal S100-Related Signature Is Correlated with Tumor Microenvironment and Predicts Prognosis in Glioma. *Dis Markers* 2022;2022:4968555.

11. Li Z, Wang J, Zhang X, et al. Proinflammatory S100A8 Induces PD-L1 Expression in Macrophages, Mediating Tumor Immune Escape. *J Immunol* 2020;204:2589-99.
12. Dey KK, Bharti R, Dey G, et al. S100A7 has an oncogenic role in oral squamous cell carcinoma by activating p38/MAPK and RAB2A signaling pathway. *Cancer Gene Ther* 2016;23:382-91.
13. Donato R, Sorci G, Giambanco I. S100A6 protein: functional roles. *Cell Mol Life Sci* 2017;74:2749-60.
14. Argyris PP, Slama ZM, Ross KF, et al. Calprotectin and the Initiation and Progression of Head and Neck Cancer. *J Dent Res* 2018;97:674-82.
15. Xiao M, Liu L, Zhang S, et al. Cancer stem cell biomarkers for head and neck squamous cell carcinoma: A bioinformatic analysis. *Oncol Rep* 2018;40:3843-51.
16. Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 2004;6:1-6.
17. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* 2015;19:A68-77.
18. Gentile A, Bensimon-Brito A, Priya R, et al. The EMT transcription factor Snai1 maintains myocardial wall integrity by repressing intermediate filament gene expression. *Elife* 2021;10:e66143.
19. Ntzifa A, Strati A, Kallergi G, et al. Gene expression in circulating tumor cells reveals a dynamic role of EMT and PD-L1 during osimertinib treatment in NSCLC patients. *Sci Rep* 2021;11:2313.
20. Shukla V, Adiga D, Jishnu PV, et al. Role of miRNA clusters in epithelial to mesenchymal transition in cancer. *Front Biosci (Elite Ed)* 2020;12:48-78.
21. Xu Z, Wang S, Ren Z, et al. An integrated analysis of prognostic and immune infiltrates for hub genes as potential survival indicators in patients with lung adenocarcinoma. *World J Surg Oncol* 2022;20:99.
22. Dennis G Jr, Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003;4:P3.
23. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2:401-4.
24. Nagy Á, Lániczky A, Menyhárt O, et al. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci Rep* 2018;8:9227.
25. Shukla S, Evans JR, Malik R, et al. Development of a RNA-Seq Based Prognostic Signature in Lung Adenocarcinoma. *J Natl Cancer Inst* 2017;109:djw200.
26. Li T, Fan J, Wang B, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res* 2016;77:e108-10.
27. Elder JT, Zhao X. Evidence for local control of gene expression in the epidermal differentiation complex. *Exp Dermatol* 2002;11:406-12.
28. Okazaki T, Honjo T. PD-1 and PD-1 ligands: from discovery to clinical application. *Int Immunol* 2007;19:813-24.
29. Yu GT, Bu LL, Huang CF, et al. PD-1 blockade attenuates immunosuppressive myeloid cells due to inhibition of CD47/SIRP α axis in HPV negative head and neck squamous cell carcinoma. *Oncotarget* 2015;6:42067-80.
30. West NR, Watson PH. S100A7 (psoriasin) is induced by the proinflammatory cytokines oncostatin-M and interleukin-6 in human breast cancer. *Oncogene* 2010;29:2083-92.
31. Liu H, Wang L, Wang X, et al. S100A7 enhances invasion of human breast cancer MDA-MB-468 cells through activation of nuclear factor- κ B signaling. *World J Surg Oncol* 2013;11:93.
32. Krop I, März A, Carlsson H, et al. A putative role for psoriasin in breast tumor progression. *Cancer Res* 2005;65:11326-34.
33. Gebhardt C, Németh J, Angel P, et al. S100A8 and S100A9 in inflammation and cancer. *Biochem Pharmacol* 2006;72:1622-31.
34. Ehrchen JM, Sunderkötter C, Foell D, et al. The endogenous Toll-like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection, autoimmunity, and cancer. *J Leukoc Biol* 2009;86:557-66.
35. Harris TM, Du P, Kawachi N, et al. Proteomic analysis of oral cavity squamous cell carcinoma specimens identifies patient outcome-associated proteins. *Arch Pathol Lab Med* 2015;139:494-507.
36. Ghavami S, Rashedi I, Dattilo BM, et al. S100A8/A9 at low concentration promotes tumor cell growth via RAGE ligation and MAP kinase-dependent pathway. *J Leukoc Biol* 2008;83:1484-92.
37. Silva EJ, Argyris PP, Zou X, et al. S100A8/A9 regulates MMP-2 expression and invasion and migration by carcinoma cells. *Int J Biochem Cell Biol* 2014;55:279-87.
38. Khammanivong A, Wang C, Sorenson BS, et al. S100A8/A9 (calprotectin) negatively regulates G2/M cell cycle progression and growth of squamous cell carcinoma. *PLoS One* 2013;8:e69395.

39. Zhu M, Wang H, Cui J, et al. Calcium-binding protein S100A14 induces differentiation and suppresses metastasis in gastric cancer. *Cell Death Dis* 2017;8:e2938.
40. Wang X, Yang J, Qian J, et al. S100A14, a mediator of epithelial-mesenchymal transition, regulates proliferation, migration and invasion of human cervical cancer cells. *Am J Cancer Res* 2015;5:1484-95.
41. Tanaka M, Ichikawa-Tomikawa N, Shishito N, et al. Co-expression of S100A14 and S100A16 correlates with a poor prognosis in human breast cancer and promotes cancer cell invasion. *BMC Cancer* 2015;15:53.
42. Meng DF, Sun R, Liu GY, et al. S100A14 suppresses metastasis of nasopharyngeal carcinoma by inhibition of NF- κ B signaling through degradation of IRAK1. *Oncogene* 2020;39:5307-22.
43. Sapkota D, Bruland O, Costea DE, et al. S100A14 regulates the invasive potential of oral squamous cell carcinoma derived cell-lines in vitro by modulating expression of matrix metalloproteinases, MMP1 and MMP9. *Eur J Cancer* 2011;47:600-10.
44. Pandey S, Osman TA, Sharma S, et al. Loss of S100A14 expression at the tumor-invading front correlates with poor differentiation and worse prognosis in oral squamous cell carcinoma. *Head Neck* 2020;42:2088-98.
45. Sapkota D, Costea DE, Blø M, et al. S100A14 inhibits proliferation of oral carcinoma derived cells through G1-arrest. *Oral Oncol* 2012;48:219-25.
46. Boye K, Maelandsmo GM. S100A4 and metastasis: a small actor playing many roles. *Am J Pathol* 2010;176:528-35.
47. Nordlinger A, Dror S, Elkahloun A, et al. Mutated MITF-E87R in Melanoma Enhances Tumor Progression via S100A4. *J Invest Dermatol* 2018;138:2216-23.
48. Stein U, Arlt F, Walther W, et al. The metastasis-associated gene S100A4 is a novel target of beta-catenin/T-cell factor signaling in colon cancer. *Gastroenterology* 2006;131:1486-500.
49. Lo JF, Yu CC, Chiou SH, et al. The epithelial-mesenchymal transition mediator S100A4 maintains cancer-initiating cells in head and neck cancers. *Cancer Res* 2011;71:1912-23.
50. Tan Y, Ma SY, Wang FQ, et al. Proteomic-based analysis for identification of potential serum biomarkers in gallbladder cancer. *Oncol Rep* 2011;26:853-9.
51. Christensen MV, Høgdall CK, Jochumsen KM, et al. Annexin A2 and cancer: A systematic review. *Int J Oncol* 2018;52:5-18.
52. Zhong J, Liu C, Chen YJ, et al. The association between S100A13 and HMGA1 in the modulation of thyroid cancer proliferation and invasion. *J Transl Med* 2016;14:80.
53. Massi D, Landriscina M, Piscazzi A, et al. S100A13 is a new angiogenic marker in human melanoma. *Mod Pathol* 2010;23:804-13.
54. Miao S, Qiu T, Zhao Y, et al. Overexpression of S100A13 protein is associated with tumor angiogenesis and poor survival in patients with early-stage non-small cell lung cancer. *Thorac Cancer* 2018;9:1136-44.
55. Wolf S, Haase-Kohn C, Pietzsch J. S100A2 in cancerogenesis: a friend or a foe? *Amino Acids* 2011;41:849-61.
56. van Dieck J, Teufel DP, Jaulent AM, et al. Posttranslational modifications affect the interaction of S100 proteins with tumor suppressor p53. *J Mol Biol* 2009;394:922-30.
57. Mueller A, Schäfer BW, Ferrari S, et al. The calcium-binding protein S100A2 interacts with p53 and modulates its transcriptional activity. *J Biol Chem* 2005;280:29186-93.
58. Cao LY, Yin Y, Li H, et al. Expression and clinical significance of S100A2 and p63 in esophageal carcinoma. *World J Gastroenterol* 2009;15:4183-8.
59. Bulk E, Sargin B, Krug U, et al. S100A2 induces metastasis in non-small cell lung cancer. *Clin Cancer Res* 2009;15:22-9.
60. Saito K, Kobayashi M, Nagashio R, et al. S100A16 is a Prognostic Marker for Lung Adenocarcinomas. *Asian Pac J Cancer Prev* 2015;16:7039-44.
61. Zhu W, Xue Y, Liang C, et al. S100A16 promotes cell proliferation and metastasis via AKT and ERK cell signaling pathways in human prostate cancer. *Tumour Biol* 2016;37:12241-50.
62. Yen MC, Huang YC, Kan JY, et al. S100B expression in breast cancer as a predictive marker for cancer metastasis. *Int J Oncol* 2018;52:433-40.

Cite this article as: Bai G, Yue S, You Y, Zhao J, Tian Z, Wang Y, Zhang Z, Yang C. An integrated bioinformatics analysis of the S100 in head and neck squamous cell carcinoma. *Transl Cancer Res* 2023;12(4):717-731. doi: 10.21037/tcr-22-1353