Received: 25 April 2022 | Revised: 1 August 2022 | Accepted: 14 August 2022

DOI: 10.1002/cam4.5170

RESEARCH ARTICLE

Cancer Medicine Open Access WILEY

COL3A1: Potential prognostic predictor for head and neck cancer based on immune-microenvironment alternative splicing

Yuchen Shen[1,2] | **Xinyu Li**[3] | **Deming Wang**[1,2] | **Liming Zhang**[1,2] | **Xiao Li**[1,2] | **Lixin Su**[1,2] | **Xindong Fan**[1,2] | **Xitao Yang**[1,2]

[1]Vascular Anomaly Center, Department of Interventional Therapy, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Shanghai Key Laboratory of Stomatology & Shanghai Research Institute of Stomatology, National Clinical Research Centre for Oral Diseases, Shanghai, China

[3]Department of Neurosurgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Correspondence
Lixin Su, Xindong Fan, and Xitao Yang, Department of Interventional Therapy, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, No. 639 Zhi Zao Ju Rd, Shanghai, 200011, China.
Email: sulixin1975@126.com; fanxindong@aliyun.com; xitao123456@126.com

Funding information
China Postdoctoral Science Foundation, Grant/Award Number: 2017M611585; Clinical Research Program of Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Grant/Award Number: JYLJ202111, JYLJ201911 and JYLJ201801; Health Clinical Research Project of Shanghai Municipal Health Commission, Grant/Award Number: 202040328; National Natural Science Foundation of China, Grant/Award Number: 81871458; Fundamental research program funding of Ninth People's Hospital affiliated to Shanghai Jiao Tong University School of Medicine, Grant/Award Number: JYZZ076

Abstract

We aimed to identify a novel prognostic biomarker for head and neck squamous cell carcinoma (HNSCC) based on tumor immunology-related alternative splicing (AS). Data for 502 HNSCC and 44 normal samples were obtained from the TCGA database and used to establish an AS-related risk model through univariate, least absolute shrinkage, and selection operator Cox regression analyses. Fresh HNSCC and normal oral tissues were surgically obtained from 44 HNSCC patients. Western blotting and quantitative reverse transcription-PCR were used to assess gene expression levels. Kaplan–Meier was performed to evaluate patients' overall survival (OS) rate. The CIBERSORT algorithm, single-sample gene set enrichment analysis, and immune checkpoint analyses were performed to compare immune activities between subgroups. The risk model was established using 10 pivotal AS events first. *Collagen Type III Alpha 1 Chain* (*COL3A1*) were screened based on $|\log2FC| \geq 1$ and FDR < 0.05 criteria. *COL3A1* expression levels in HNSCC tissues were elevated relative to normal tissues ($p < 0.001$). Moreover, *COL3A1* was a reliable biomarker for HNSCC patients' prognostic prediction in both cohorts ($p < 0.001$, $p = 0.0085$, respectively). COL3A1 protein ($p = 0.0054$) and mRNA ($p < 0.0001$) levels were correlated with HNSCC differentiation. Furthermore, the T stage was correlated with *COL3A1* expression ($p = 0.043$), and *COL3A1* expression was an independent prognostic predictor for HNSCC patients ($p = 0.006$). Compared with the risk model, COL3A1 was better at

Yuchen Shen, Xinyu Li, Deming Wang contributed equally to this work.

evaluating immune cell infiltrations, immune activities, and immune checkpoint gene expressions of HNSCC lesions.

KEYWORDS
alternative splicing, COL3A1, HNSCC, immune checkpoint, prognosis

1 | INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is the common malignancy in the head and neck region, accounting for more than 880,000 patients/year, 7.5 billion people, and more than 450,000 deaths/year, 7.5 billion people.[1] HNSCC patients' 5-year survival rate remains low (50–55%).[2]

Despite high recurrence rates and low cancer survival rates, treatment advances have been made. Cancer immunotherapy, which depends on T cell–mediated immune responses, is the most significant among these advances. For instance, adoptive T cell therapy based on CART and TCR-T has shown promising therapeutic effects in cancer patients, making it a promising strategy for cancer management.[3] Therefore, immunotherapy is a potential approach for overcoming the limitations associated with traditional cancer therapeutic strategies.

In 2019, Li et al. presented a systematic review of immunotherapy regarding alternative mRNA splicing.[4] Processing of pre-mRNA transcripts is an essential step in the final function of gene products.[5] Most human genes contain many exons, and contiguous intron sequences must be linked to pre-transcriptional mRNA to form the mature product.[6] AS, breaking down pre-mRNA into a single mature transcript, contributes to the diversity of the transcriptome and proteome.[7] Alternate acceptor site (AA), alternate donor site (AD), alternate promoter (AP), alternate terminator (AT), exon skip (ES), mutually exclusive exons (ME), and retained intron (RI) are main types.[8] Polypeptides of tumor-specific ribonucleic acid splicing events may bind to MHC class I (MHC-I) molecules, which serve as a novel epitope.[9–11] Additional analysis of the Clinical Proteomics Oncology Consortium (CPTAC) 63 breast and ovarian cancers showed that 68% of the tumors contained one or more alternatively spliced neoepitopes. In contrast, only 30% of tumors contained neoepitopes derived from somatic events with single-nucleotide variants.[4] These findings form the basis for studies on immune-related AS events in tumors.

Despite several promising risk models of HNSCC based on AS,[12–14] deficiencies in these models exist, and the correlation between AS and tumor immunology in HNSCC has yet to be elucidated. Therefore, a systematic analysis of HNSCC-related AS events should be performed from a tumor immune microenvironment (TIM) perspective. In this study, we aimed at i. establishing a risk model for HNSCC patients based on pivotal AS events; ii. screening genes can be used as prognostic predictors for HNSCC patients; and iii. we are evaluating the accuracy of the risk-model and prognostic gene according to the status of the TIM.

2 | MATERIALS AND METHODS

2.1 | Data acquisition for alternative splicing events

RNA sequencing data for 502 HNSCC and 44 normal samples were retrieved from The Cancer Genome Atlas (TCGA) and the HNSCC cohort (https://tcga-data.nci.nih.gov/tcga/). SpliceSeq was used to assess mRNA splicing patterns of samples. Percent Spliced In (PSI) values, ranging from zero to one, which is commonly used in quantifying AS events, were imputed for seven different AS events.[15] AS events with PSI values >75% were included. After screening HNSCC patients, we finally included 487 HNSCC samples and 44 normal controls. All clinicopathological data were obtained from TCGA. Immunohistochemical (IHC) data were downloaded from Proteinatlas (http://www.proteinatlas.org/pathology) to assess the expressions of differentially expressed genes in HNSCC.

2.2 | Patients and sample collection

Forty-four patients' fresh HNSCC and their healthy tissues were collected, all of which were resected during surgery in the Department of Oral and Maxillofacial-Head and Neck Oncology, Shanghai Ninth People's Hospital from April 2010 to October 2014.[16] The patients including 25 males and 19 females, aging from 35 to 86, with an average of 59 ± 12.4. For tissue collection, the method was the same as our previous study.[17] Subsequently, the pathological diagnosis was carried out in the Department of Pathology, Shanghai Ninth Public Hospital. The clinical evaluation of all cases was carried out according to the 8th edition of the TNM

classification of head and neck cancer.[18] Based on pathological reports, among the 44 cases, 15 cases were characterized as high-differentiation HNSCC, 20 cases were characterized as moderate-differentiation HNSCC, and 9 cases were characterized as low-differentiation HNSCC. A telephone survey followed up all 44 HNSCC patients until October 2019.

2.3 | Identification of survival-related alternative splicing events and construction of the prognostic model

We performed a univariate Cox regression analysis to identify AS events related to overall survival (OS) with $p < 0.05$. Upset plots were generated using UpsetR for quantitative analysis of interaction sets among the seven types of OS-related AS. We selected the highest AS events associated with survival as candidates for each splice type to fit Least absolute shrinkage and selection operator (LASSO) Cox analysis for HNSCC. A risk score was calculated for each prognostic model from the sum of the PSI values of the identified AS events and the corresponding coefficients derived from multivariate Cox analysis. The formula for calculating the risk score was as previously described[17]:

$$\text{Risk score} = \Sigma_i^n \beta_i \times PSI_i$$

N represents the quantity of AS events, β are the coefficients, while PSI denotes PSI values for specific AS events.

2.4 | Identification of differentially expressed genes

Before comparisons, we normalized the expression data in the TCGA database as fragment per kilobase million (FPKM) values. The "limma" package was used to identify differentially expressed genes (DEGs) using $|\log2FC| \geq 1$ and FDR < 0.05 as the cut-offs. Moreover, based on each DEG expression level, we equally divided the TCGA cohort into low- and high-expression subgroups.

2.5 | Immune function analyses

Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) database (https://bioinformatics.mdanderson.org/public-software/estimate/) was used to generate immune score and stromal score, which represent the degrees of infiltrations of immune and stromal cells in tumor tissues, respectively. Data on tumor-infiltrating immune cells were estimated using the CIBERSORT algorithm. The "gsva" package was used to perform a single-sample gene set enrichment analysis (ssGSEA) and to calculate the infiltrating immune cells' scores and responses to assess the activities of immune pathways.

2.6 | Protein extraction and western immunoblotting

For western immunoblotting (WB), we referred to our previous study for the relevant operating steps.[2] Primary antibodies against COL3A1 (ab184993, 1:1000 dilution; Abcam [UK]), GAPDH (WL01547, 1:1500 dilution; Wanlei Bio), and the secondary goat-anti-rabbit antibody (ZB-2301; Zhong Shan Golden Bridge) were used in this study. WB bands were detected by Alpha FluorChem FC3 (Protein Simple), whereas imaging data were quantified by the ImageJ software as previously mentioned.[2]

2.7 | Total RNA extraction, reverse transcription, and qRT-PCR

Total RNA was extracted from clinical specimens using RNAiso Plus (9109, TAKARA), after which cDNA was synthesized from 10 μg of the extracted total mRNA using the PrimeScript™RT Master Mix (RR036A, TAKARA), according to the manufacturer's instructions. Quantitative reverse transcription-PCR (qRT-PCR) was performed using Power SYBR Green PCR Master Mix (4,367,659, Thermo Fisher Scientific) and Q2000B (LongGene). Primers for COL3A1 and GAPDH were obtained with reference to PrimerBank (https://pga.mgh.harvard.edu/primerbank/index.html) (Table S1). Relative differences in expression between samples were analyzed by the $2^{-\Delta\Delta Ct}$ method. GAPDH was selected as an internal reference.

2.8 | Immune checkpoint analysis based on the risk model

We screened and extracted 46 immune checkpoint genes (ICGs) from previous studies[19–22] (Table S2). Then, gene expression data for the TCGA cohort was used to analyze the correlation between risk score and each ICG. Second, the data was used to assess the expression profiles of ICGs between subgroups. Both processes were performed using R.

2.9 | Statistical analysis

R (v4.0.2), SPSS 24.0, and Prism Graphpad 8 were used for statistical analysis. Two-tailed students' t-tests and one-way ANOVA were used to analyze the correlation between groups. The Kaplan–Meier analysis evaluated the OS rate, and the two-tailed log-rank test and Mann–Whitney U-test were used to compare the differences between the two groups. Univariate and multivariate Cox regression analyses were carried out to calculate the survival hazard ratio (HR) and 95% confidence interval (CI). p value <0.05 was considered statistically significant.

3 | RESULTS

3.1 | Overview of alternative splicing event profiles and construction of the prognostic risk model in head and neck squamous cell carcinoma

First, UpSet plot were generated to visualize interactive sets between the seven types of AS events. Figure 1A,B showed that a total of 42,849 AS events were detected from 10,123 genes. Univariate Cox analysis found 2893 survival-related AS events within 1733 genes in our HNSCC cohort (all $p<0.05$). For each AS type, the top 20 survival-related events in HNSCC are shown in Figure S1 (only 8 for ME). Then, we performed a LASSO Cox analysis for HNSCC, and 10 pivotal AS events for establishing the risk model were screened (Figure 1C,D, Table S3). After multivariate Cox analysis had been performed, our risk-score formula was as follows:

$$\begin{aligned}\text{Risk score} = &\left(2.19^{*}PSI_{\text{GPR56|36575|AP}}\right)+\left(1.80^{*}PSI_{\text{KCNAB2|367|RI}}\right)+\left(-0.78^{*}PSI_{\text{E2F3|75492|AP}}\right)+\\&\left(-1.13^{*}PSI_{\text{OSBPL3|79025|AD}}\right)+\left(0.91^{*}\,PSI_{\text{ISLR|31677|AP}}\right)+\left(1.70^{*}\,PSI_{\text{SFR1|13036|AP}}\right)+\\&\left(-1.92^{*}PSI_{\text{LIPG|45487|AT}}\right)+\left(2.23^{*}PSI_{\text{ATP9B|46234|ES}}\right)+\left(1.24^{*}PSI_{\text{BAIAP2|44097|RI}}\right)+\left(-1.93^{*}PSI_{\text{COL3A1|485142|ES}}\right).\end{aligned}$$

HNSCC patients were equally divided into low- and high-risk subgroups based on the median score (Figure 1E). Compared with patients in the low-risk group, patients in the high-risk group exhibited higher death rates and shorter survival times (Figure 1F). The heat map presents the overall situation of the 10 fundamental AS events, as shown in Figure 1G.

Kaplan–Meier analysis showed that the patients in the low-risk group enjoyed better prognostic outcomes relative to their counterparts ($p<0.001$, Figure 1H). Univariate and multivariate analyses suggested that the risk score could be an independent factor influencing HNSCC patients' prognoses ($p<0.001$, Figure S2A). The ROC curves showed that the prediction model based on risk scores had better differentiating abilities (AUC for risk score = 0.763, Figure S2B; AUC at 1 year = 0.763, AUC at 2 years = 0.756; AUC at 3 years = 0.745, Figure S2C).

3.2 | Comparisons of immune activities between the risk-subgroups

Then, we used the ESTIMATE database to assess the status of each HNSCC patient's TIM. As shown in Figure 2A, patients from the low-risk group received higher immune scores compared with the high-risk group ($p<0.0001$); however, no significant differences were found between the two subgroups with regard to stromal scores ($p = 0.31$, Figure 2B). These findings indicate that low-risk patients had higher infiltration rates of immune cells and more active immune responses in the tumor microenvironment than the high-risk patients. For immune cells, low-risk patients' resting $CD4^{+}$ T cells, M0, and M2 macrophages were significantly increased compared with high-risk patients (all $p<0.001$, Figure 2C). With regard to immune score, low-risk patients exhibited more infiltration of macrophages, mast cells and type II IFN responses, relative to high-risk patients ($p<0.001$, $p<0.01$, $p<0.01$, Figure 2D,E).

3.3 | Identification of differentially expressed gene and immune-related analysis

Next, we screened the TCGA cohort's DEGs from the risk model's 10 pivotal AS events. Only *Collagen Type III Alpha 1 Chain* (*COL3A1*) met the $|\log2FC|\geq1$ and FDR <0.05 criteria (data not shown). HNSCC patients from the TCGA cohort exhibited significant differences in *COL3A1* expression levels compared with normal patients ($p<0.001$, Figure 3A). Then, we equally assigned the HNSCC patients into two subgroups according to their *COL3A1* levels. Notably, Kaplan–Meier analysis of OS based on *COL3A1* expression levels indicated that *COL3A1* could be used for prognostic prediction of HNSCC patients. HNSCC patients with suppressed *COL3A1* expression levels were associated with high OS rates ($p<0.001$, Figure 3B). Since AS can affect HNSCC patients' TIM, then, based on the

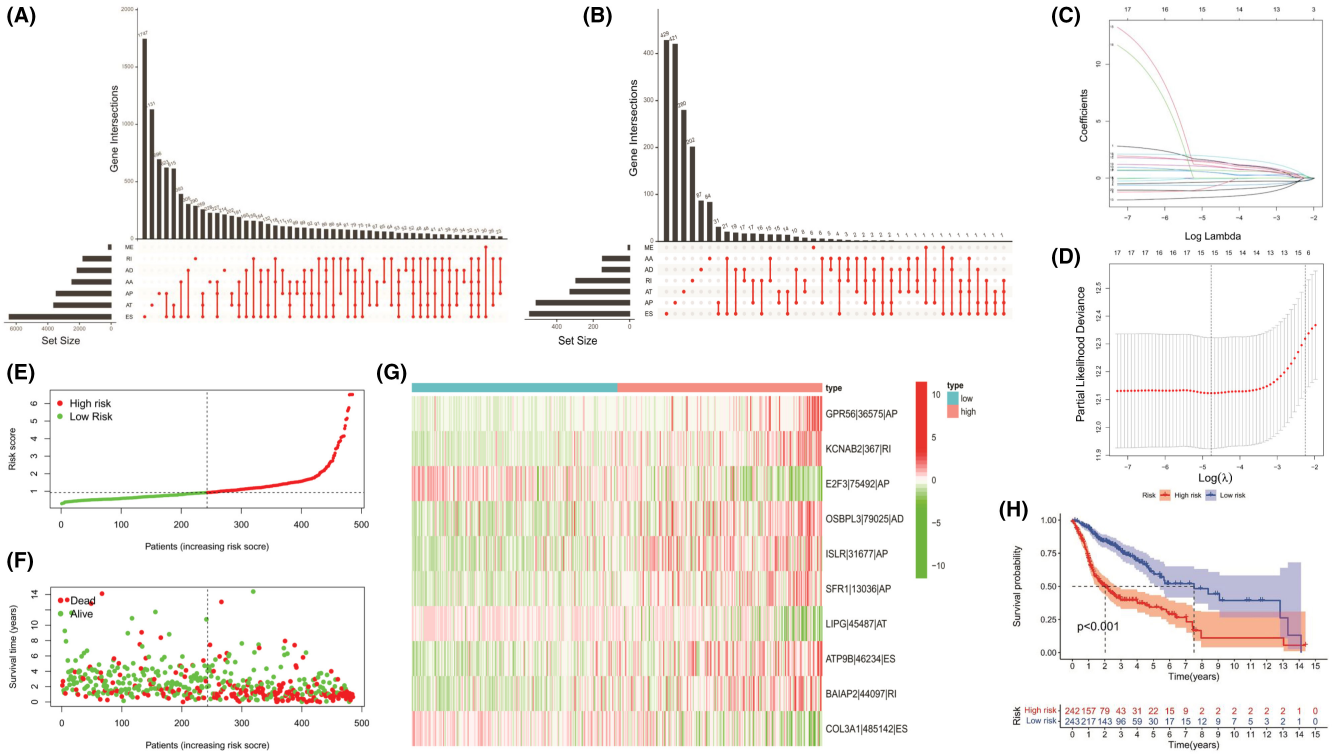

FIGURE 1 Overall review of the risk model's alternative splicing (AS) events and construction processes. (A) Upset plot of interaction among seven standard AS events in head and neck squamous cell carcinoma (HNSCC). (B) Upset plot of interactions among the seven survival-associated AS events in HNSCC. (C) LASSO regression of survival-associated AS events. (D) Cross-validation to optimize parameter selection in LASSO regression. (E) Distribution of patients based on the risk score. (F) All patients' survival status (low-risk population: dotted line on the left side; high-risk population: dotted line on the right side). (G) Heatmap (green: low PSI value; red: high PSI value) of the 10 pivotal AS events between the low-risk group (brilliant blue) and the high-risk group (red). (H) Kaplan–Meier curve between the low-risk group and the high-risk group.

ESTIMATE database, we evaluated whether *COL3A1* has the same effects. Interestingly, the group with low *COL3A1* expression had a higher immune score ($p < 0.0001$) but not stromal score ($p = 0.075$), relative to the high-expression group, suggesting that HNSCC patients with low COL3A1 levels had more activated immune status in TIM than their counterparts (Figure 3C,D). With regard to immune cells, patients with elevated COL3A1 levels had more resting $CD4^+$ T cells and resting natural killer (NK) cells but less activated $CD4^+$ T cells compared with *COL3A1* low-expressing patients (all $p < 0.001$, Figure 3E). Patients with elevated COL3A1 levels were generally associated with less immune activity than those with suppressed levels (Figure 3F).

3.4 | Correlations between *COL3A1* expressions and survival outcomes for head and neck squamous cell carcinoma patients

To analyze the expression levels of *COL3A1* in HNSCC, first, we assessed overall differences between paraffin-embedded HNSCC samples and normal oral samples using the IHC data in Proteinatlas. Protein levels of COL3A1 in HNSCC samples were higher than those of the normal oral mucosa samples (Figure 4A). Next, we performed WB to assess protein levels of COL3A1 in fresh tissues of practical HNSCC patients. Figure 4B,C showed that COL3A1 levels were about threefold higher in fresh HNSCC compared with normal oral tissues ($p < 0.0001$). Interestingly, protein levels of COL3A1 significantly increased as histopathologic grades for HNSCC decreased (Figure 4D, $p < 0.01$).

Next, we extracted total mRNA from sections of HNSCC and normal control tissues and assessed mRNA levels of *COL3A1*. *COL3A1* mRNA expressions were elevated in HNSCC tissues compared with normal control tissues (Figure 4E, $p < 0.001$). Surprisingly, as shown in Figure 4F, though the general tendency of *COL3A1* mRNA levels among each HNSCC differentiated tissue was close to WB results, *COL3A1* mRNA levels in poorly differentiated HNSCC tissues were nearly 50-fold compared with normal tissues. Then, we equally assigned the 44 HNSCC patients into two subgroups according to their *COL3A1* mRNA levels and performed Kaplan–Meier analyses to compare OS rates between the two groups. As shown in Figure 4G, the

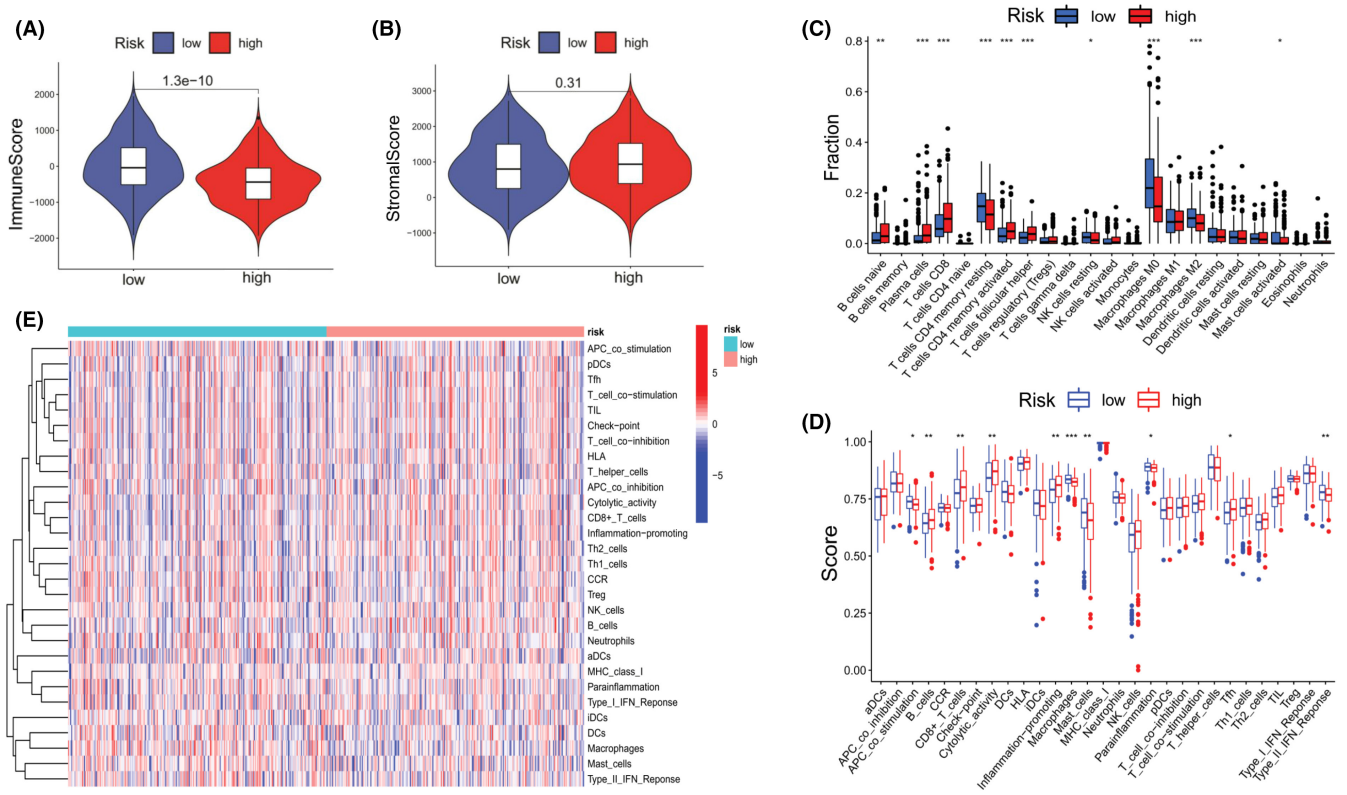

FIGURE 2 Immune activities of the low- and high-risk groups. (A) Immune scores for both subgroups. (B) Stromal scores for both subgroups. (C) Immune cell infiltrations in both subgroups. (D) Scores of immune activities for both subgroups. (E) Overall heatmap of immune cells and responses (red: high expression; blue: low expression) of pyroptosis genes between the low-risk (brilliant blue) and high-risk (red) groups. (Two-tailed students' *t*-test, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

cumulative five-year OS rate was 40.9% (18 of 44). Kaplan–Meier analysis demonstrated that OS outcomes for patients with elevated *COL3A1* levels were significantly low relative to those with suppressed *COL3A1* levels ($p = 0.0085$).

3.5 | Correlations between *COL3A1* levels and head and neck squamous cell carcinoma patients' clinicopathological features

Correlations between *COL3A1* expression levels and potential risk factors (smoking and alcohol histories) for 44 HNSCC patients and the association between *COL3A1* levels, sex, and age (divided by 60 years) were analyzed. Among the 44 HNSCC patients, 15 cases were high differentiation, 20 were moderate differentiation, and 9 were low differentiation. These patients were classified into the high *COL3A1* expressing group (22 cases) and the low *COL3A1* expressing group (22 cases). As Table 1 showed, expression levels of *COL3A1* were significantly correlated with the T stage ($p = 0.043$). *COL3A1* expression levels were not significantly different from clinical parameters such as sex, age, smoking status, alcohol history, clinical status, and lymph node metastasis.

3.6 | Univariate and multivariate analysis of head and neck squamous cell carcinoma patients' prognoses

Univariate Cox models were used to estimate individual clinical parameters for OS to establish the potential prognostic value. As shown in Table 2, significant factors for OS were T stage ($p < 0.01$), clinical tumor stage ($p = 0.048 < 0.05$) and *COL3A1* expression levels ($p < 0.01$). Subsequently, these factors were included in the multivariate Cox regression model, which further showed that *COL3A1* expression was the independent factor for OS ($p = 0.006$, HR = 3.731, 95% CI: 1.458–9.546).

3.7 | Correlation between immune checkpoint genes and the prognostic model or *COL3A1* expression

Given the important roles of immune checkpoints in tumor immune escape, we determined differences in expressions for each of the ICGs between low- and high-risk groups. These ICGs include TNFRSF18, TMIGD2, CD27, PD1, IDO2, CTLA4, ADORA2A, NRP1,

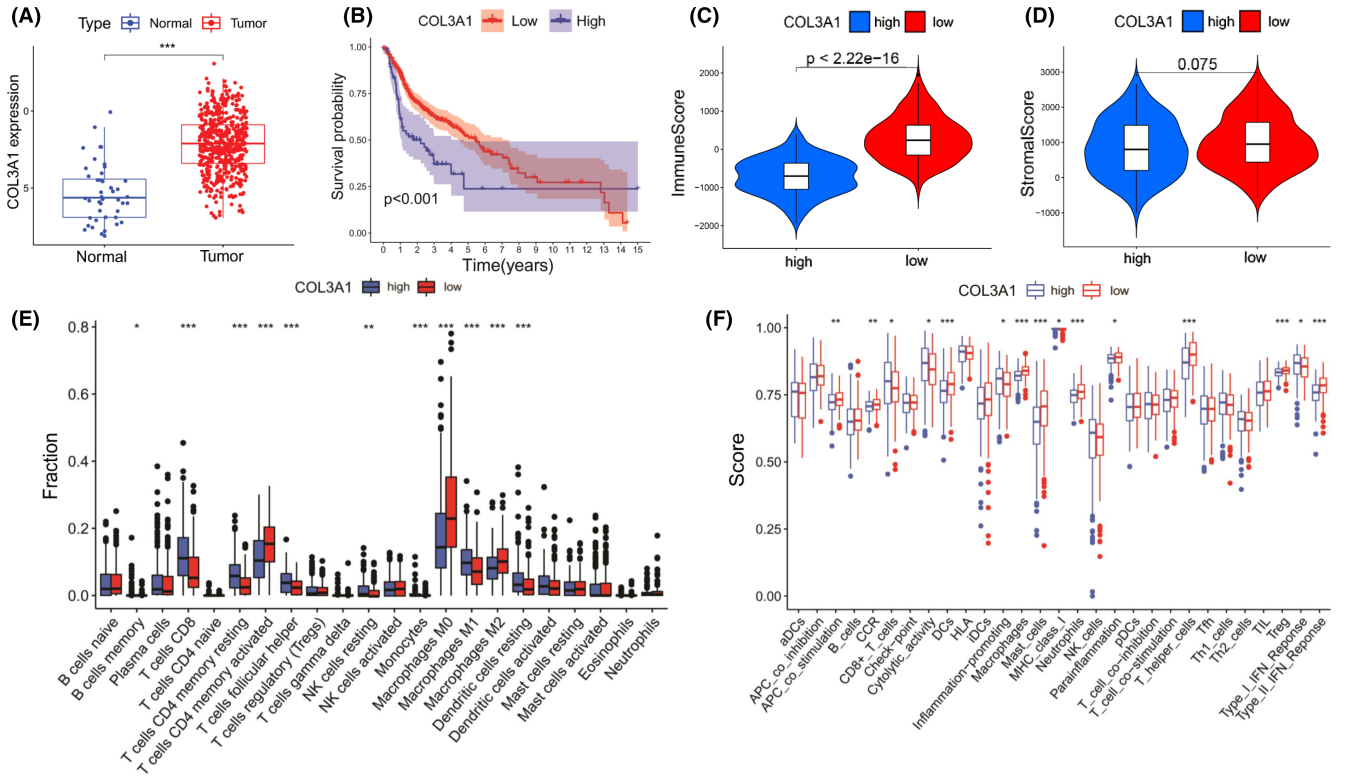

FIGURE 3 Prognosis and immune status between head and neck squamous cell carcinoma (HNSCC) patients with different *COL3A1* expression levels. (A) *COL3A1* expression levels between HNSCC and normal tissues. (B) Kaplan–Meier analysis between *COL3A1* low- and high-expression groups. (C) Immune scores for both subgroups. (D) Stromal scores for both subgroups. (E) Immune cell infiltrations in both subgroups. (F) Scores of immune activities for both subgroups. (Two-tailed students' *t*-test, $*p<0.05$; $**p<0.01$; $***p<0.001$).

TNFRSF25, CD276, TNFRSF4, TNFRSF14, and CD44 exhibited the highest differences between the two subgroups (Figure 5A, all $p<0.001$). Next, we evaluated the tendency of gene expression between ICGs and the *COL3A1*. As shown in Figure 5B, patients from *COL3A1* low- and high-expression groups were endowed with significant differences in gene expression status of ICGs. Between the subgroups, *TNFSF4*, *LAIR1*, *CD200*, *CD28*, *CD80*, *TNFRSF9*, *NRP1*, *TNFRSF25*, *CD26*, *PD1LG2*, *CD86*, and *HAVCR2* were differentially expressed (all $p<0.001$). Furthermore, we found that superfamilies of TNF and their receptors, namely TNFSF and TNFRSF, were dominant among ICGs. The expression of *TNFSF4*, *TNFSF14*, *TNFRSF18*, and *TNFRSF25* in HNSCC patients and their correlation with COL3A1 were detailed in Figure 5C-F (all $p<0.05$).

4 | DISCUSSION

AS is a common gene expression regulation mechanism that allows genes to produce several different types of mRNA. AS is tightly regulated in different tissues, cell types, and stages of differentiation.[23–25] In many malignancies, AS is essential in tumorigenesis.[26–29] Kahles et al. found that healthy tissue bears fewer AS events than tumors.[9] Moreover, the impact of AS on the immune system is a potential predictor for responses to immune checkpoint-related therapy. Cascades of immune checkpoints, including those controlled by programmed cell death 1 (PD-1) or T-cell cytotoxic antigen 4 (CTLA-4), could be treated as negative regulators of immunity. Therefore, suppressing these cascades with antibodies has transformed many cancers' treatment, leading to high levels of enduring tumor responses.[30–32] In HNSCC, studies should evaluate how AS can benefit prevention, diagnosis, treatment, and prognostic prediction from a tumor immune perspective.

In this study, we first established a prognostic risk model based on AS events in HNSCC patients. Ten events were screened and included in the risk formula. Even though our risk model was shown to have the ability for prognostic prediction, there were some limitations when assessing HNSCC patients' immune statuses. For instance, as shown in Figure 2C, low-risk patients exhibited high infiltration levels of resting $CD4^+$ T cells and resting NK cells but less activated $CD4^+$ T cells relative to high-risk patients. These findings are inconsistent with natural immune microenvironment conditions for patients with high-grade HNSCC.[33,34] Therefore, the risk

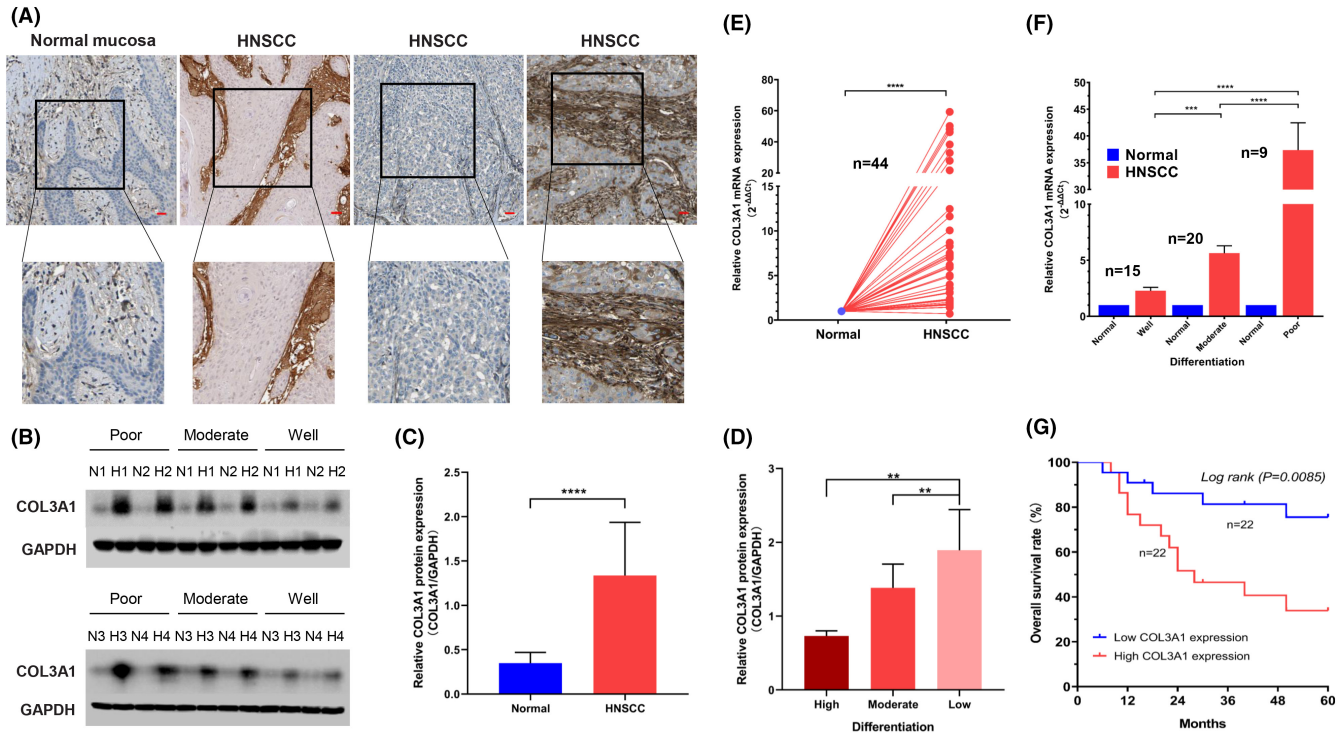

FIGURE 4 Comparisons of *COL3A1* expressions between head and neck squamous cell carcinoma (HNSCC) and normal tissues and survival outcomes. (A) Representative immunohistochemical images of paraffin-embedded HNSCC and normal oral samples (Bar: 50 μm). (B) Representative WB images for four patients from each HNSCC differentiation. (C) Relative COL3A1 protein levels in HNSCC and normal tissues as evaluated by band intensities and areas in WB ($n = 44$). (D) Comparisons of COL3A1 protein levels among differentially differentiated HNSCC tissues ($p = 0.0054$). (E) Comparisons of relative *COL3A1* mRNA levels between 44 HNSCC and 44 normal control tissues. (F) Comparisons of relative *COL3A1* mRNA levels among each HNSCC differentiation ($p < 0.0001$). (G) Kaplan–Meier curves for 44 HNSCC patients based on their *COL3A1* mRNA levels. (One-way ANOVA and two-tailed students' *t*-test, $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$).

model must refine the key genes to enhance its predictive accuracy.

Collagen Type III Alpha 1 Chain, also known as Collagen Alpha-1(III) Chain, is a protein encoded by the *COL3A1* gene in humans. Type III collagen is involved in many cellular functions despite its interaction with integrins and cell surface receptors.[35] As shown in Figure 3E,F, regarding the TIM of HNSCC, COL3A1 exhibited better performance than the risk model (Figures C and D). Even though the risk model and *COL3A1* could be used to predict HNSCC patients' prognoses, the latter exhibited a better ability to predict HNSCC patients' tumor immune responses.

COL3A1 has been shown to promote neoplasia in several types of malignancies[36,37]; however, its role in HNSCC evolution has not been conclusively determined. Our study demonstrated that COL3A1 exhibited the potency to become a prognostic biomarker for HNSCC, for it was more activated in cancer tissues than normal tissues, not only in the TCGA cohort but also in the practical patient cohort. Elevated *COL3A1* levels were associated with poor prognostic outcomes for HNSCC patients, and as the differentiation of HNSCC tissues improved, COL3A1 expression levels reduced. For patients' case history and clinicopathological characteristics, only the T stage was correlated with COL3A1 (Table 1). Through univariate and multivariate analyses, COL3A1 expression was an independent factor in predicting the HNSCC patient's OS rates (Table 2).

The present study found that COL3A1|485,142|ES is a critical AS event as it exhibited a relatively high absolute coefficient value in the formula. Per a former report, the most frequent pattern of abnormal splicing of COL3A1 is exon skipping (ES).[38] COL3A1 plays a pivotal role in the extracellular matrix. It has been reported that the abnormal expression of *COL3A1* was closely associated with fibrosis in the cardiovascular system and spinal ligaments, indicating COL3A1's capacity to maintain the functions of fibroblasts.[39,40] Rachel et al. found that patients with vascular Ehlers-Danlos syndrome, a genetic disease caused by the *COL3A1* exon skipping mutations and presented with increased vascular fragility, suffered from systemic inflammation.

Given the above findings, we evaluated the underlying immune mechanisms. Cancer cells can prevent immune

TABLE 1 Correlation between the expression of COL3A1 and patients' case history and clinicopathological characteristics

Characteristics	Number of cases (%)	COL3A1 expression		p^a value
		High (%)	Low (%)	
Sex				0.373
Male	25 (56.8)	11 (25.0)	14 (31.8)	
Female	19 (43.2)	11 (25.0)	8 (18.2)	
Age				0.373
≤60	25 (56.8)	14 (31.8)	11 (25.0)	
>60	19 (43.2)	8 (18.2)	11 (25.0)	
Smoking history				0.757
Present	29 (65.9)	14 (31.8)	15 (34.1)	
Absent	15 (34.1)	8 (18.2)	7 (15.9)	
Alcohol history				0.764
Present	27 (61.4)	13 (29.5)	14 (31.8)	
Absent	17 (38.6)	9 (20.5)	8 (18.2)	
T stage				**0.043***
T1 + T2	32 (72.7)	13 (29.5)	19 (43.2)	
T3 + T4	12 (27.3)	9 (20.5)	3 (6.8)	
Clinical stage				0.062
I-II	28 (63.6)	11 (25.0)	17 (38.6)	
III-IV	16 (36.4)	11 (25.0)	5 (11.4)	
Lymphatic metastasis				0.365
Positive	17 (38.6)	10 (22.7)	7 (15.9)	
Negative	27 (61.4)	12 (27.3)	15 (34.1)	

Bold value indicate Statistical significance.

[a]All statistical tests were two-sided.

*$p < 0.05$.

surveillance and progression through a series of mechanisms, including activation of immune checkpoint pathways to suppress immune responses against cancer.[41] Tumor immunotherapy based on immune-checkpoint blockade is a novel strategy for coping with refractory neoplasms. By blocking negative regulators of the innate immunity such as cytotoxic T lymphocyte antigen 4 (CTLA-4) and programmed cell death 1 (PD-1) or their ligands, programmed cell death ligand 1 (PD-1), immune checkpoint blockade enhances antitumor immunity.[42] Due to the potential effects of AS events on TIM via affecting ICGs, we evaluated the correlation between ICGs and our risk model or *COL3A1* expression. As shown in Figure 5A,B, high-risk group patients exhibited more ICGs activities, implying that cancer cells in high-risk patients' HNSCC lesions have an enhanced ability to escape supervisory control and the threat from the immune system. These outcomes were also found in HNSCC patients whose *COL3A1* expression levels were higher, which could explain why both high-risk HNSCC patients and high-*COL3A1* expression patients are prone to poor prognostic outcomes.

Another interesting finding was that genes of the *TNFSF* and *TNFRSF* were differentially expressed between the high- and low-risk subgroups and high- and low-*COL3A1* expression groups (Figure 5A,B). The above has revealed the potential role of *TNFSF* and *TNFRSF* members in facilitating HNSCC progression and, therefore, as potential targets for HNSCC immunotherapy. First, *TNFRSF25* was differentially expressed between low- and high-*COL3A1* expression subgroups (Figure 5E). *TNFRSF25 is* mainly expressed in activated T cells, antigen-presenting cells (APCs), and phagocytes.[43–45] Activation of TNFRSF25 by its ligand, TNFSF15, is important for T-cell proliferation during viral infection control as it enhances interferon-g (IFN-g) synthesis.[46] Co-stimulatory *TNFRSF25* promotes the differentiation of Th1 and Th9 $CD4^+$ T cells. It dampens the suppressive capacity of T_{reg} cells, implying that an agonistic antibody for TNFRSF25 is an attractive agent for cancer immunotherapy.[47,48] In contrast to *COL3A1* expression, the risk model did not represent the actual immue situation of HNSCC lesions as high-risk patients exhibited elevated *TNFRSF25* levels relative to low-risk patients.

TABLE 2 Univariate and multivariate analyses of clinical characteristics for overall survival in patients with head and neck squamous cell carcinoma

Characteristics	Amount (%)	Univariate		Multivariate	
		p^a value	HR[b] (95% CI[c])	p^a value	HR[b] (95% CI[c])
Age					
≤60	25 (56.8)	0.321	1.580 (0.612–4.078)		
>60	19 (43.2)				
Sex					
Male	25 (56.8)	0.706	0.835 (0.329–2.123)		
Female	19 (43.2)				
Smoking history					
Present	29 (65.9)	0.102	2.111 (0.750–5.942)		
Absent	15 (34.1)				
Alcohol history					
Present	27 (61.4)	0.228	0.540 (0.212–1.378)		
Absent	17 (38.6)				
T stage					
T1 and T2	32 (72.7)	**<0.01****	3.609 (1.177–11.07)	0.073	
T3 and T4	12 (27.3)				
Clinical stage					
Stage I and II	28 (63.6)	**0.048***	2.450 (0.922–6.506)	0.798	
Stage III and IV	16 (36.4)				
Lymphatic metastasis					
Positive	17 (38.6)	0.105	0.478 (0.184–1.241)		
Negative	27 (61.4)				
COL3A1 expression					
Low	22 (50)	**<0.01****	1.275 (1.118–1.358)	**0.006****	3.731 (1.458–9.546)
High	22 (50)				

Bold values indicate Statistical significance.

[a]All statistical tests were two-sided.

[b]HR, Hazard ratio.

[c]CI, confidence interval.

*$p < 0.05$; **$p < 0.01$.

A similar advantage of COL3A1 was also confirmed by *TNFRSF18* expression (Figure 5F). *TNFRSF18* is mainly expressed in APCs and medullary thymic epithelial cells. Furthermore, *TNFRSF18* is highly expressed in T_{reg} progenitors, and activating ligands may promote its development in mature thymic T_{reg} cells by upregulating CD25.[49,50] The binding of TNFRSF18 to its ligand, TNFSF18, can suppress T_{reg} cell recruitment, weaken their inhibitory functions and activate the mitogen-activated protein kinase (MAPK)/ extracellular regulated protein kinases (ERK) pathway as well as nuclear factor kappa B (NF-κB) signaling, thereby promoting T-cell proliferation, secretion of pro-inflammatory cytokines, and enhancing antitumor functions.[51] Previous findings from studies involving the agonistic TNFRSF18 antibody, TRX518, laid the groundwork for clinical trials in malignant melanoma or other solid cancers.[52]

Finally, it was also noteworthy that patients with elevated COL3A1 levels had a higher level of resting $CD4^+$T cells and resting NK cells, but fewer activated CD4+ T cells than their counterparts (Figure 3E,F). Differences in *TNFSF* and *TNFRSF* gene expressions might shed light on these intriguing results. According to the previous studies, *TNFRSF18* is weakly expressed by naive $CD4^+$T cells but is upregulated on activation. In comparison, *TNFRSF25* is upregulated in activated $CD4^+$T cells, B cells, NK cells, and NKT cells.[53] Thus, our results in Figure 5E,F consolidated the standpoint that HNSCC patients with higher *COL3A1* expression generally possessed a relatively suppressed condition of immune cells owing to low-level *TNFRSF18* and *TNFRSF25*.

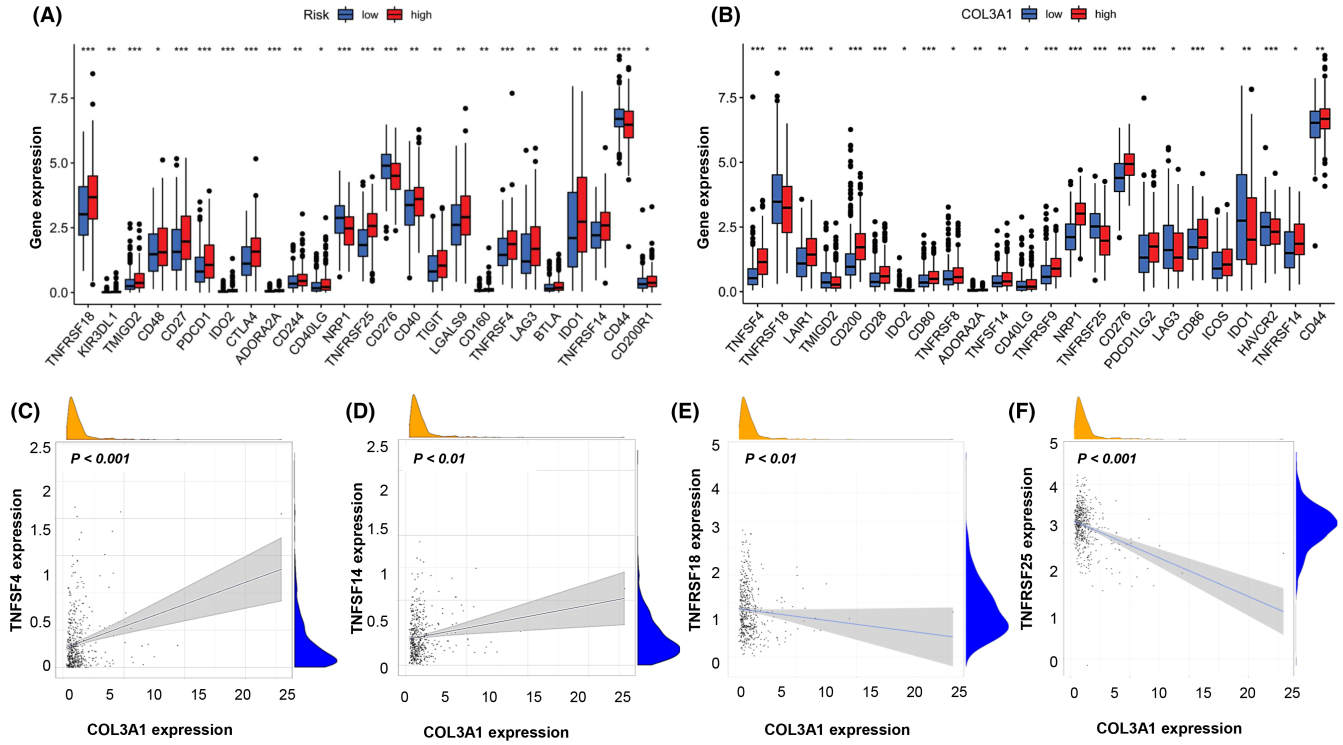

FIGURE 5 Correlation between immune checkpoint genes (ICGs) and *COL3A1*. (A) Differences in ICGs levels between low- and high-risk groups. (B) Differences in ICG levels between *COL3A1* low- and high-expression groups. (B-F) The correlations between gene expression levels of *TNFSF4*, *TNFSF14*, *TNFRSF18*, *TNFRSF25*, and *COL3A1* expression. (Two-tailed students' *t*-test, $*p < 0.05$; $**p < 0.01$; $***p < 0.001$).

Although we found that COL3A1 has the potential for prognostic prediction of HNSCC, a total of 44 HNSCC patients and their control samples represent relatively small sample size. Additionally, immune cells can intimately interact with stromal cells and further impact tumor progression, invasion, and metastasis. Apart from immune score, it should be noted that there was no significant discrepancy in the stromal score between the subgroups either in the risk model or in COL3A1 expression (Figures 2B and 3D). This phenomenon reminded us that the risk model and COL3A1 alone have fallen short of presenting the real stromal status in the tumor microenvironment. Despite these limitations, our findings will inform on HNSCC treatment.

AUTHOR CONTRIBUTIONS

XTY: Conceptualization, Supervision, Writing – review & editing; YCS: Investigation, Writing – original draft; XYL: Data curation, Formal analysis; XDF: Funding acquisition; DMW: Methodology; LXS: Project administration; LMZ: Validation; XL: Visualization.

FUNDING INFORMATION

This study was funded by the National Natural Science Foundation of China (No. 81871458), the Health Clinical Research Project of Shanghai Municipal Health Commission (No. 202040328), Clinical Research Program of Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine (No. JYLJ201801, JYLJ201911, JYLJ202111) and the China Postdoctoral Science Foundation (No. 2017M611585), Fundamental research program funding of Ninth People's Hospital affiliated to Shanghai Jiao Tong University School of Medicine (No. JYZZ076).

CONFLICT OF INTEREST

The authors declared that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

This study was approved by the Human Research Ethics Committee of the Ninth People's Hospital, Shanghai JiaoTong University School of Medicine (Shanghai, China).

PATIENT CONSENT STATEMENT

Given the retrospective nature of this study, informed consent was waived.

ORCID

Yuchen Shen https://orcid.org/0000-0002-1181-5198
Xitao Yang https://orcid.org/0000-0001-6031-0412

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249.
2. Shen Y, Zhang L, Piao S, et al. NUDT1: A potential independent predictor for the prognosis of patients with oral squamous cell carcinoma. *J Oral Pathol Med*. 2020;49(3):210-218.
3. Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: From T cell basic science to clinical practice. *Nat Rev Immunol*. 2020;20(11):651-668.
4. Frankiw L, Baltimore D, Li G. Alternative mRNA splicing in cancer immunotherapy. *Nat Rev Immunol*. 2019;19(11):675-687.
5. Liu Y, Gonzalez-Porta M, Santos S, et al. Impact of alternative splicing on the human proteome. *Cell Rep*. 2017;20(5):1229-1241.
6. Weatheritt RJ, Sterne-Weiler T, Blencowe BJ. The ribosome-engaged landscape of alternative splicing. *Nat Struct Mol Biol*. 2016;23(12):1117-1123.
7. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. *Nature*. 2010;463(7280):457-463.
8. Zhang Y, Yan L, Zeng J, et al. Pan-cancer analysis of clinical relevance of alternative splicing events in 31 human cancers. *Oncogene*. 2019;38(40):6678-6695.
9. Kahles A, Lehmann KV, Toussaint NC, et al. Comprehensive analysis of alternative splicing across tumors from 8,705 patients. *Cancer Cell*. 2018;34(2):211-24 e6.
10. Jayasinghe RG, Cao S, Gao Q, et al. Systematic analysis of splice-site-creating mutations in cancer. *Cell Rep*. 2018;23(1):270, e3-281.
11. Smart AC, Margolis CA, Pimentel H, et al. Intron retention is a source of neoepitopes in cancer. *Nat Biotechnol*. 2018;36(11):1056-1058.
12. Liang Y, Song J, He D, et al. Systematic analysis of survival-associated alternative splicing signatures uncovers prognostic predictors for head and neck cancer. *J Cell Physiol*. 2019;234:15836-15846.
13. Zhang S, Wu X, Diao P, et al. Identification of a prognostic alternative splicing signature in oral squamous cell carcinoma. *J Cell Physiol*. 2020;235(5):4804-4813.
14. Cao R, Zhang J, Jiang L, et al. Comprehensive analysis of prognostic alternative splicing signatures in Oral squamous cell carcinoma. *Front Oncol*. 2020;10:1740.
15. Ryan MC, Cleland J, Kim R, Wong WC, Weinstein JN. SpliceSeq: A resource for analysis and visualization of RNA-seq data on alternative splicing and its functional impacts. *Bioinformatics*. 2012;28(18):2385-2387.
16. Zhang LM, Su LX, Hu JZ, et al. Epigenetic regulation of VENTXP1 suppresses tumor proliferation via miR-205-5p/ANKRD2/NF-kB signaling in head and neck squamous cell carcinoma. *Cell Death Dis*. 2020;11(10):838.
17. Shen Y, Li X, Wang D, et al. Novel prognostic model established for patients with head and neck squamous cell carcinoma based on pyroptosis-related genes. *Transl Oncol*. 2021;14(12):101233.
18. Huang SH, O'Sullivan B. Overview of the 8th edition TNM classification for head and neck cancer. *Curr Treat Options Oncol*. 2017;18(7):40.
19. Zeng D, Li M, Zhou R, et al. Tumor microenvironment characterization in gastric cancer identifies prognostic and Immunotherapeutically relevant gene signatures. *Cancer Immunol Res*. 2019;7(5):737-750.
20. Ravi R, Noonan KA, Pham V, et al. Bifunctional immune checkpoint-targeted antibody-ligand traps that simultaneously disable TGFbeta enhance the efficacy of cancer immunotherapy. *Nat Commun*. 2018;9(1):741.
21. Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer*. 2016;16(5):275-287.
22. Hu FF, Liu CJ, Liu LL, Zhang Q, Guo AY. Expression profile of immune checkpoint genes and their roles in predicting immunotherapy response. *Brief Bioinform*. 2021;22(3):bbaa176.
23. Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol*. 2017;18(7):437-451.
24. Wong JJ, Ritchie W, Ebner OA, et al. Orchestrated intron retention regulates normal granulocyte differentiation. *Cell*. 2013;154(3):583-595.
25. Pimentel H, Parra M, Gee S, et al. A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis. *Nucleic Acids Res*. 2014;42(6):4031-4042.
26. Ellis MJ, Ding L, Shen D, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature*. 2012;486(7403):353-360.
27. Biankin AV, Waddell N, Kassahn KS, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature*. 2012;491(7424):399-405.
28. Harbour JW, Roberson ED, Anbunathan H, Onken MD, Worley LA, Bowcock AM. Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. *Nat Genet*. 2013;45(2):133-135.
29. Imielinski M, Berger AH, Hammerman PS, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell*. 2012;150(6):1107-1120.
30. Freeman GJ, Long AJ, Iwai Y, et al. Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J Exp Med*. 2000;192(7):1027-1034.
31. Chambers CA, Kuhns MS, Egen JG, Allison JP. CTLA-4-mediated inhibition in regulation of T cell responses: mechanisms and manipulation in tumor immunotherapy. *Annu Rev Immunol*. 2001;19:565-594.
32. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science*. 2018;359(6382):1350-1355.
33. Wu L, Mao L, Liu JF, et al. Blockade of TIGIT/CD155 signaling reverses T-cell exhaustion and enhances antitumor capability in head and neck squamous cell carcinoma. *Cancer Immunol Res*. 2019;7(10):1700-1713.
34. Cillo AR, Kurten CHL, Tabib T, et al. Immune landscape of viral- and carcinogen-driven head and neck cancer. *Immunity*. 2020;52(1):183-99 e9.
35. Kuivaniemi H, Tromp G. Type III collagen (COL3A1): Gene and protein structure, tissue distribution, and associated diseases. *Gene*. 2019;707:151-171.

36. Gao YF, Zhu T, Chen J, Liu L, Ouyang R. Knockdown of collagen alpha-1(III) inhibits glioma cell proliferation and migration and is regulated by miR128-3p. *Oncol Lett*. 2018;16(2):1917-1923.
37. Yuan L, Shu B, Chen L, et al. Overexpression of COL3A1 confers a poor prognosis in human bladder cancer identified by co-expression analysis. *Oncotarget*. 2017;8(41):70508-70520.
38. Schwarze U, Goldstein JA, Byers PH. Splicing defects in the COL3A1 gene: Marked preference for 5′ (donor) splice-site mutations in patients with exon-skipping mutations and Ehlers-Danlos syndrome type IV. *Am J Hum Genet*. 1997;61(6):1276-1286.
39. Bartoli F, Evans EL, Blythe NM, et al. Global PIEZO1 gain-of-function mutation causes cardiac hypertrophy and fibrosis in mice. *Cell*. 2022;11(7):1199.
40. Hayashi F, Morimoto M, Higashino K, et al. Myofibroblasts are increased in the dorsal layer of the hypertrophic ligamentum flavum in lumbar spinal canal stenosis. *Spine J*. 2022;22(4):697-704.
41. Darvin P, Toor SM, Sasidharan Nair V, Elkord E. Immune checkpoint inhibitors: Recent progress and potential biomarkers. *Exp Mol Med*. 2018;50(12):1-11.
42. Postow MA, Sidlow R, Hellmann MD. Immune-related adverse events associated with immune checkpoint blockade. *N Engl J Med*. 2018;378(2):158-168.
43. Fang L, Adkins B, Deyev V, Podack ER. Essential role of TNF receptor superfamily 25 (TNFRSF25) in the development of allergic lung inflammation. *J Exp Med*. 2008;205(5):1037-1048.
44. Shih DQ, Kwan LY, Chavez V, et al. Microbial induction of inflammatory bowel disease associated gene TL1A (TNFSF15) in antigen presenting cells. *Eur J Immunol*. 2009;39(11):3239-3250.
45. Longman RS, Diehl GE, Victorio DA, et al. CX(3)CR1(+) mononuclear phagocytes support colitis-associated innate lymphoid cell production of IL-22. *J Exp Med*. 2014;211(8):1571-1583.
46. Twohig JP, Marsden M, Cuff SM, et al. The death receptor 3/TL1A pathway is essential for efficient development of antiviral CD4(+) and CD8(+) T-cell immunity. *FASEB J*. 2012;26(8):3575-3586.
47. Richard AC, Tan C, Hawley ET, et al. The TNF-family ligand TL1A and its receptor DR3 promote T cell-mediated allergic immunopathology by enhancing differentiation and pathogenicity of IL-9-producing T cells. *J Immunol*. 2015;194(8):3567-3582.
48. Meylan F, Song YJ, Fuss I, et al. The TNF-family cytokine TL1A drives IL-13-dependent small intestinal inflammation. *Mucosal Immunol*. 2011;4(2):172-185.
49. Clouthier DL, Watts TH. Cell-specific and context-dependent effects of GITR in cancer, autoimmunity, and infection. *Cytokine Growth Factor Rev*. 2014;25(2):91-106.
50. Mahmud SA, Manlove LS, Schmitz HM, et al. Costimulation via the tumor-necrosis factor receptor superfamily couples TCR signal strength to the thymic differentiation of regulatory T cells. *Nat Immunol*. 2014;15(5):473-481.
51. Mei Z, Huang J, Qiao B, Lam AK. Immune checkpoint pathways in immunotherapy for head and neck squamous cell carcinoma. *Int J Oral Sci*. 2020;12(1):16.
52. Ward-Kavanagh LK, Lin WW, Sedy JR, Ware CF. The TNF receptor superfamily in co-stimulating and co-inhibitory responses. *Immunity*. 2016;44(5):1005-1019.
53. Dostert C, Grusdat M, Letellier E, Brenner D. The TNF family of ligands and receptors: Communication modules in the immune system and beyond. *Physiol Rev*. 2019;99(1):115-160.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Shen Y, Li X, Wang D, et al. COL3A1: Potential prognostic predictor for head and neck cancer based on immune-microenvironment alternative splicing. *Cancer Med*. 2023;12:4882-4894. doi: 10.1002/cam4.5170