



Research Article

Identification of potential biomarkers associated with immune infiltration in the esophageal carcinoma tumor microenvironment

 Zhicheng Wang^{1,2}, Meilin Chen^{1,2}, Yanbing Qiu^{1,2}, Yuqin Yang^{1,2}, Yumei Huang^{1,2}, Xiaoxu Li^{1,2} and  Wenling Zhang^{1,2}

¹Department of Laboratory Medicine, The Third Xiangya Hospital, Central South University, Changsha, Hunan, People's Republic of China; ²Department of Laboratory Medicine, Xiangya School of Medicine, Central South University, Changsha, Hunan, People's Republic of China

Correspondence: Wenling Zhang (zhangwenling73@126.com)



Tumor immune cell infiltration was significantly correlated with the progression and the effect of immunotherapy in cancers including esophageal carcinoma (ESCA). However, no biomarkers were identified which were associated with immune infiltration in ESCA. In the present study, a total of 128 common differentially expressed genes (DEGs) were identified between esophageal squamous cell carcinomas (ESCC) and esophageal adenocarcinomas (EAC). The results of gene ontology (GO) enrichment and Reactome pathway analysis displayed that the up-regulated DEGs were mainly involved in the regulation of extracellular matrix (ECM), while the down-regulated DEGs were mainly involved in the regulation of cornification and keratinocyte differentiation. The most significant module of up-regulated DEGs was selected by Molecular Complex Detection (MCODE). Top ten similar genes of COL1A2 were explored, then validation and the prognostic analysis of these genes displayed that COL1A2, COL1A1, COL3A1, ZNF469 and Periostin (POSTN) had the prognostic value which were up-regulated in ESCA. The expressions of COL1A2 and its four similar genes were mainly correlated with infiltrating levels of macrophages and dendritic cells (DCs) and showed strong correlations with diverse immune marker sets in ESCA. To summarize, COL1A2 and its four similar genes were identified as the potential biomarkers associated with immune infiltration in ESCA. These genes might be applied to immunotherapy for ESCA.

Introduction

As a kind of gastrointestinal cancer, esophageal carcinoma (ESCA) ranks seventh according to incidence (572000 new cases) and sixth in mortality totally (509000 deaths) in 2018 [1]. The progression and chemoresistance of ESCA are important factors that impact patient survival. For those patients with locally advanced ESCA, the 5-year survival rate is approximately 20%. The standard treatment includes chemotherapy or chemoradiotherapy with surgery [2]. In recent years, tumor immunologic therapies have also shown promising early results in ESCA. And tumor immune cell infiltration significantly correlates with the progression and the effect of immunotherapy in cancers [3,4]. However, the specific molecular mechanism about immunotherapy for ESCA yet remains unclear and some genes associated with immune infiltration are required to be found.

With the development of gene chip technology and sequencing technology, bioinformatics is becoming more and more important [5–7]. Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) are two open databases, which provide gene chip data and sequencing data about the cancers. A lot of studies were published for analyzing the data from GEO and TCGA in ESCA, including the profiles of long noncoding RNA (lncRNA), microRNA (miRNA), and messenger RNA (mRNA). Interleukin 8

Received: 01 July 2020
Revised: 16 December 2020
Accepted: 24 December 2020

Accepted Manuscript online:
05 February 2021
Version of Record published:
17 February 2021

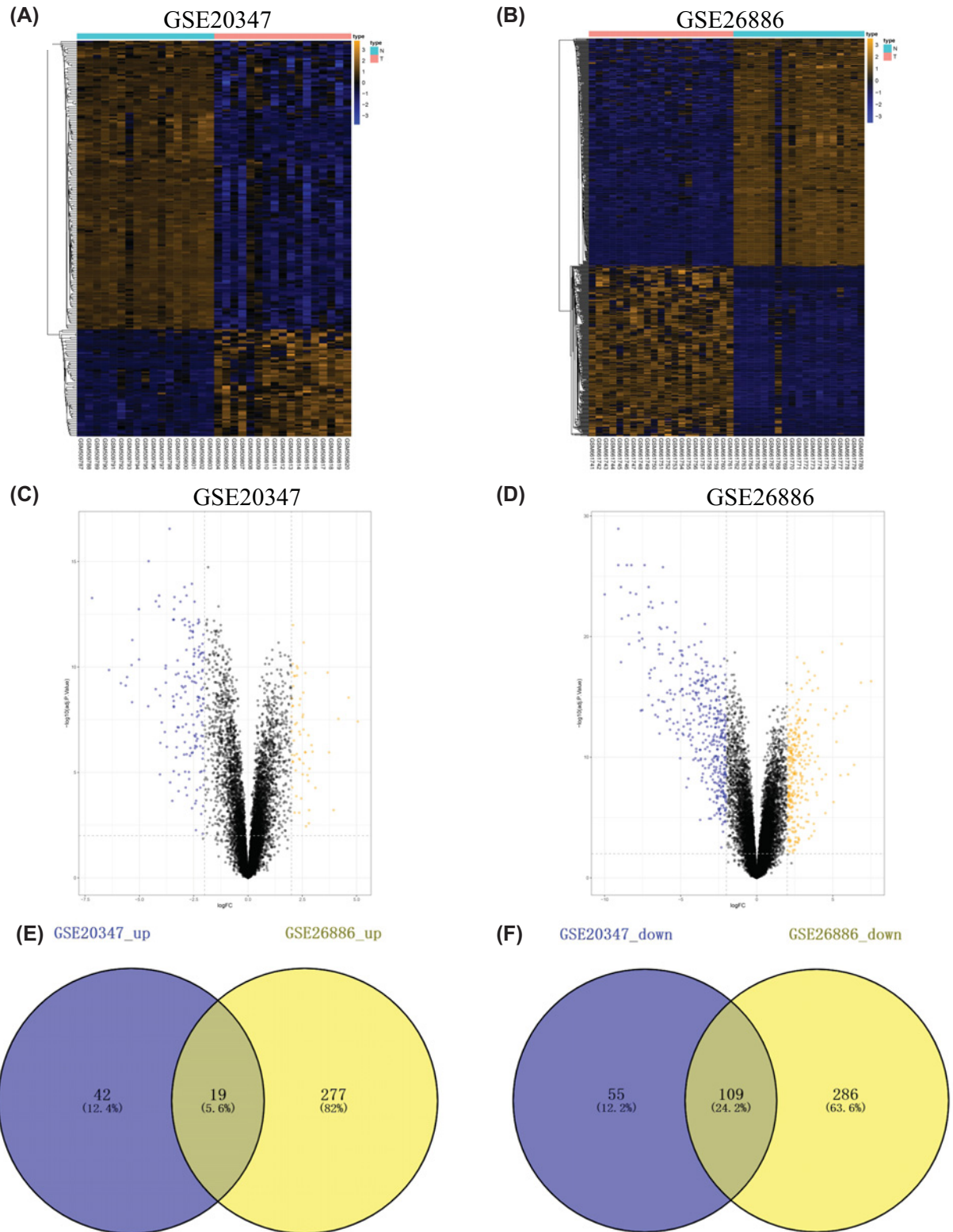


Figure 1. Identification of DEGs in GSE20347 and GSE26886

Heatmaps of DEGs in (A) GSE20347 and (B) GSE26886 were displayed. Volcano plots of DEGs in (C) GSE20347 and (D) GSE26886 were displayed. Orange: up-regulated DEGs; Blue: down-regulated DEGs; N: normal; T: tumor. (E) Up-regulated DEGs and (F) down-regulated DEGs were integrated between GSE20347 and GSE26886 by using the Venn diagram.

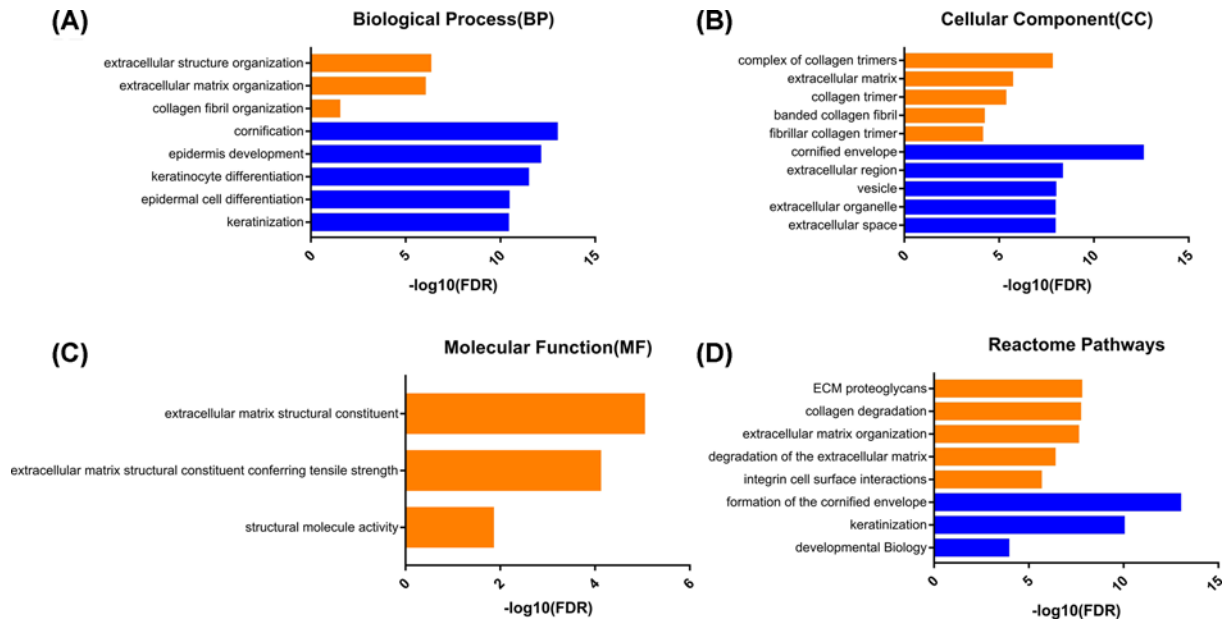


Figure 2. GO and Reactome pathway analysis of DEGs

(A) BP, (B) CC, (C) MF, and (D) Reactome pathway enrichment analysis were displayed by the bar chart. The bars of up-regulated DEGs enrichment results were marked in orange and the bars of down-regulated DEGs enrichment results were marked in blue.

(IL8) and C–X–C chemokine receptor type 7 (CXCR-7) were predicted to differentially express in ESCA tissues compared with normal esophageal tissues and significantly correlate to the prognosis of patients with ESCA [8]. Liu et al. constructed the network of lncRNA, miRNA, and mRNA and found eight lncRNAs including WDFY3-AS2, CASC8, UGDH-AS1, RAP2C-AS1, AC007128.1, AC016205.1, AC092803.2, and AC079949.2 correlated with overall survival of the patients with ESCA [9]. Zeng et al. built the regulatory networks of miRNA-TF and miRNA-gene and identified the top five miRNAs (miRNA-93, miRNA-21, miRNA-4746, miRNA-196a-1, and miRNA-196a-2) with diagnostic value [10]. However, the immunotherapy-related genes were certainly not identified in these studies in ESCA.

In the present study, the common differentially expressed genes (DEGs) were screened between esophageal squamous cell carcinomas (ESCC) and esophageal adenocarcinoma (EAC) by analyzing the datasets from GEO. Then the gene ontology (GO) analysis, Reactome pathway analysis, and protein–protein interaction (PPI) analysis of common DEGs were performed. Subsequently, COL1A2 with its similar genes were selected in Gene Expression Profiling Interactive Analysis (GEPIA) and validated in UALCAN, meanwhile the prognostic significance analysis was performed. Finally, the immune infiltration analysis was carried out to explain the immunological significance of COL1A2 with its similar genes. The findings in this report revealed the important role of COL1A2 with its similar genes in regulating immune infiltration and suggested that these genes might be immunotherapy-related genes in ESCA.

Materials and methods

Selection of microarray data from GEO

Two microarray datasets (GSE20347 and GSE26886) were downloaded from GEO. GSE20347 was based on the GPL571 platform (Affymetrix Human Genome U133A 2.0 Array) which contained 34 samples, including 17 pairs of tumors and their matching normal adjacent tissues from ESCC patients. GSE26886 was based on the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array) which contained 40 samples, including 21 EAC samples and 19 normal esophageal squamous epithelium samples.

Screening of common DEGs between ESCC and EAC

We downloaded the raw CEL data and standardized them by the Affy package of R language [11]. Then the differential gene analysis was carried out using the limma R package and genes with an adjusted *P*-value <0.01 and |log₂ fold change (FC)| > 2 were considered DEGs [12]. The heatmaps and volcano plots of DEGs were displayed. The common DEGs of two datasets were integrated using an online web tool, Venny 2.1, and the common DEGs including up-regulated and down-regulated genes were saved for subsequent analysis.

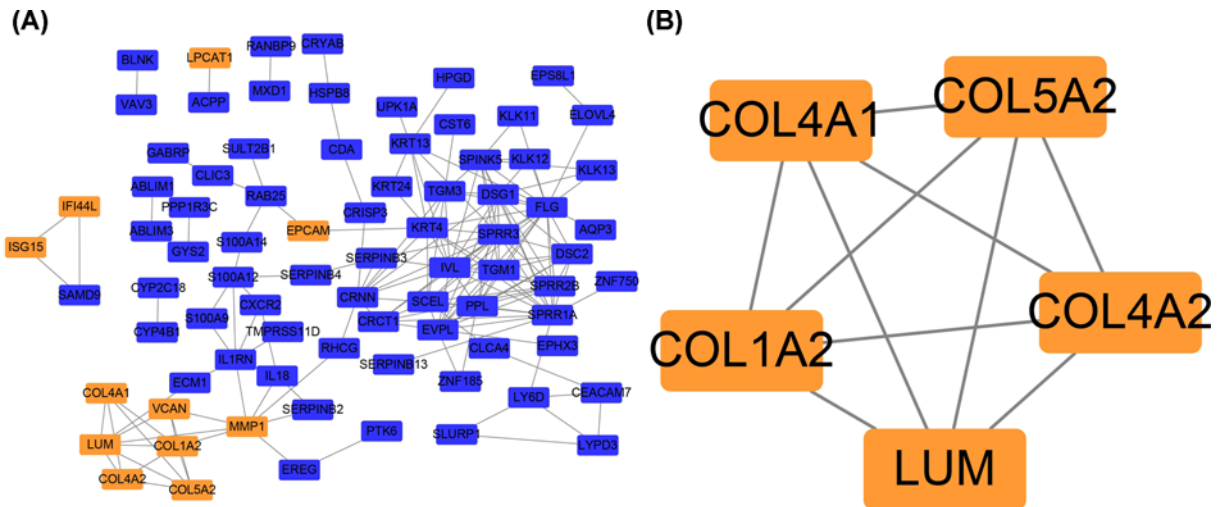


Figure 3. Construction of PPI network of DEGs and identification of the most significant module of up-regulated DEGs
(A) PPI network of DEGs. The up-regulated genes were marked in orange and the down-regulated genes were marked in blue. (B) The most significant module of up-regulated DEGs. The up-regulated genes were marked in orange.

GO and Reactome pathway analysis of common DEGs

The GO PANTHER version 14.0 was used to perform GO functional and Reactome pathway analysis of the common DEGs [13]. And the GO functional analysis of integrated DEGs refers to three parts: biological process (BP), cell component (CC), and molecular function (MF). FDR < 0.05 was considered to have a statistically significant difference.

PPI network and module analysis

We constructed a PPI network of the common DEGs using the Search Tool for the Retrieval of Interacting Genes (STRING), an online web tool for predicting the correlation of protein function [14]. And the PPI network was visualized by Cytoscape software [15]. Then, the most significant module of up-regulated DEGs was selected by Molecular Complex Detection (MCODE), which was a plug-in application of Cytoscape [16].

Validation of COL1A2 and its similar genes in UALCAN

UALCAN can analyze cancer OMICS data which is a comprehensive, user-friendly, and interactive web tool [17]. The expressions of COL1A2 and its similar genes were validated in primary ESCA samples compared with normal esophageal squamous epithelium samples. $P < 0.05$ was considered to indicate a statistically significant difference.

Similar genes and prognostic significance analysis in GEPIA

GEPIA is a web-based tool used to deliver fast and customizable functionalities based on TCGA and GTEx data. Its primary functions included differential expression analysis, patient survival analysis, similar gene detection, and so on [18]. Similar genes and prognostic significance analysis were executed here to study our target genes and $P < 0.05$ was considered to have a statistically significant difference.

Immune infiltration analysis in Tumor Immune Estimation Resource

Tumor Immune Estimation Resource (TIMER) is an interactive web tool to comprehensively investigate the molecular characterization of tumor-immune interactions. Six major functional modules allow users to interactively explore the correlations between immune infiltrates and a wide range of factors, including gene expression, somatic mutations, and somatic copy number alterations, and so on [19]. The correlations between the expression levels of COL1A2 with its similar genes and the infiltration levels of various immune cells were analyzed in ESCA, including B cells, CD8⁺ T cells, CD4⁺ T cells, neutrophils, macrophages, and dendritic cells (DCs). And we also analyzed the relationships between the expression levels of COL1A2 with its similar genes and the immune marker sets including CD8⁺ T cell, T cell (general), B cell, monocytes, tumor-associated macrophages (TAMs), M1 macrophages, M2 macrophages,

Table 1 The similar genes of COL1A2 explored by GEPIA and ranked by Pearson correlation coefficient (PCC) in ESCA tumors

| Genes symbol | Gene ID | PCC |
|--------------|--------------------|------|
| COL1A1 | ENSG00000108821.13 | 0.98 |
| COL3A1 | ENSG00000168542.12 | 0.97 |
| SPARC | ENSG00000113140.10 | 0.93 |
| ZNF469 | ENSG00000225614.2 | 0.92 |
| POSTN | ENSG00000133110.14 | 0.92 |
| CTSK | ENSG00000143387.12 | 0.92 |
| COL5A1 | ENSG00000130635.15 | 0.91 |
| LRRC15 | ENSG00000172061.8 | 0.91 |
| ADAM12 | ENSG00000148848.14 | 0.90 |

neutrophils, natural killer (NK) cells, DCs, T-helper (Th) cells, T-helper 17 (Th17) cells, follicular helper T (Tfh) cells, regulatory T cells (Tregs) and exhausted T cells based on previous studies [20–22].

Results

Identification of common DEGs between ESCC and EAC

To find the common DEGs from the two subtypes of ESCA, we analyzed the ESCC chip expression dataset GSE20347 and EAC chip expression dataset GSE26886 (adjusted $P < 0.01$ and $|\log_2 FC| > 2$). The GSE20347 dataset contained 225 DEGs, including 61 up-regulated genes and 164 down-regulated genes (Figure 1A,C). The GSE26886 dataset contained 691 DEGs, including 296 up-regulated genes and 395 down-regulated genes (Figure 1B,D). The common up-regulated DEGs and down-regulated DEGs from the two kinds of datasets were identified respectively, which were displayed by the Venn diagram, including 19 up-regulated genes and 109 down-regulated genes (Figure 1E,F).

GO and reactome pathway analysis of common DEGs

DEGs GO and Reactome pathway analysis were performed with PANTHER. The data shown in the figure needed to meet the following criteria: (1) the $FDR < 0.05$; (2) the top five significant enrichment results. We can find the BP of up-regulated genes which were mainly enriched in the extracellular structure organization, extracellular matrix (ECM) organization, and collagen fibril organization. The down-regulated genes were mainly enriched in the cornification, epidermis development, keratinocyte differentiation, epidermal cell differentiation, and keratinization (Figure 2A). And the CC of up-regulated genes were mainly enriched in the complex of collagen trimers, ECM, collagen trimer, banded collagen fibril, and fibrillar collagen trimer. The down-regulated genes were mainly enriched in the cornified envelope, extracellular region, vesicle, extracellular organelle, and extracellular space (Figure 2B). The MF of up-regulated genes were mainly enriched in the ECM structural constituent, ECM structural constituent conferring tensile strength and structural molecule activity, but the results of down-regulated genes were not displayed because the $FDR > 0.05$ (Figure 2C). As for the pathway analysis, the up-regulated DEGs were significantly enriched in the ECM proteoglycans, collagen degradation, ECM organization, degradation of the ECM, and integrin cell surface interactions, while the down-regulated DEGs were significantly enriched in the formation of the cornified envelope, keratinization, and developmental biology (Figure 2D).

Construction of PPI network and module analysis

The PPI network of DEGs was constructed by using the Cytoscape, which contained 81 nodes and 174 edges (Figure 3A). The most significant module of up-regulated DEGs was identified by using MCODE, which contained COL1A2, COL5A2, COL4A2, LUM, and COL4A1 (Figure 3B).

Prognostic significance analysis and validation of COL1A2 and its similar genes

COL1A2 was a subtype of Type I collagen, the pieces of evidence suggest its mRNA expression in ESCC tumors was higher than that in normal tissues. And it was associated with the survival of ESCC patients [23]. In the following study, we selected COL1A2 for our important target gene. In the first place, we explored the top ten similar genes of COL1A2 in ESCA by using GEPIA (Table 1). In order to validate our results, we analyzed the differential expression of COL1A2 and its similar genes between 11 normal esophageal squamous epithelium samples and 184 primary ESCA

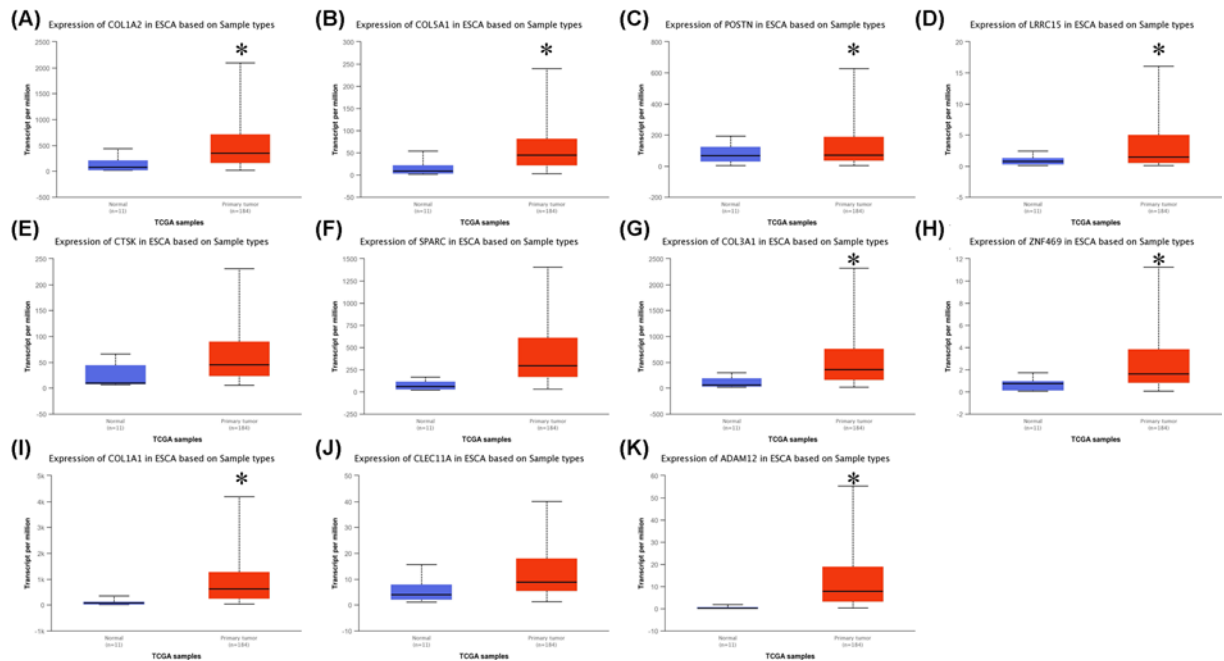


Figure 4. Boxplots of mRNA expression levels of COL1A2 and its similar genes in normal esophageal squamous epithelium samples and primary ESCA samples
(A) COL1A2 (B) COL5A1 (C) POSTN (D) LRRC15 (E) CTSK (F) SPARC (G) COL3A1 (H) ZNF469 (I) COL1A1 (J) CLEC11A (K) ADAM12. Normal, normal esophageal squamous epithelium samples; Primary tumor, primary ESCA samples. *Normal vs Primary tumor and $P < 0.05$. $P < 0.05$ was considered statistically significant.

samples by using UALCAN (Figure 4A–K). Except for CTSK, SPARC, and CLEC11A, all of the genes expressed higher in primary ESCA samples compared with normal esophageal squamous epithelium samples. Meanwhile, we investigated whether these genes contributed to the prognosis of ESCA patients, GEPIA was used to analyze the disease-free survival (DFS) of COL1A2 and its similar genes in patients. As shown in Figure 5A–K, the higher expression of COL1A2 and its similar genes showed worse DFS in ESCA patients, except for COL5A1, LRRC15, ADAM12 and CLEC11A. These genes that have been validated and had prognostic value will be used for later study.

The expression levels of COL1A2 and its similar genes correlate with the infiltration levels of immune cells in ESCA

The infiltration of immune cells such as TAMs and Tregs in the microenvironment can promote esophageal carcinogenesis [24]. Therefore, we explored the relationship between the expression of COL1A2 with its similar genes and the infiltrating immune cells in ESCA using the TIMER database. Tumor purity plays an important role in the analysis of immune cell infiltration about clinical tumor samples by genomic approaches [25]. After adjusting the analysis by tumor purity, we found that the expression levels of COL1A2, COL1A1, COL3A1, ZNF469 and Periostin (POSTN) were significantly correlated with tumor purity in ESCA. And the expression levels of these genes were positively correlated with the infiltration levels of macrophages and DCs, as well as the expression levels of ZNF469 was also positively correlated with the infiltration levels of CD4⁺ T cells (Figure 6).

Correlation analysis between mRNA levels of COL1A2 with its similar genes and immune marker sets

To investigate the relationship between COL1A2 with its similar genes and the different immune infiltrating cells, the correlations between COL1A2 with its similar genes and immune marker sets of various immune cells of ESCA were explored in the TIMER database. We analyzed the correlations between the expression of COL1A2 with its similar genes and immune marker genes of diverse immune cells, including CD8⁺ T cells, T cells (general), B cells, monocytes, TAMs, M1 and M2 macrophages, neutrophils, NK cells, and DCs in ESCA. Besides, the diverse T cells, for

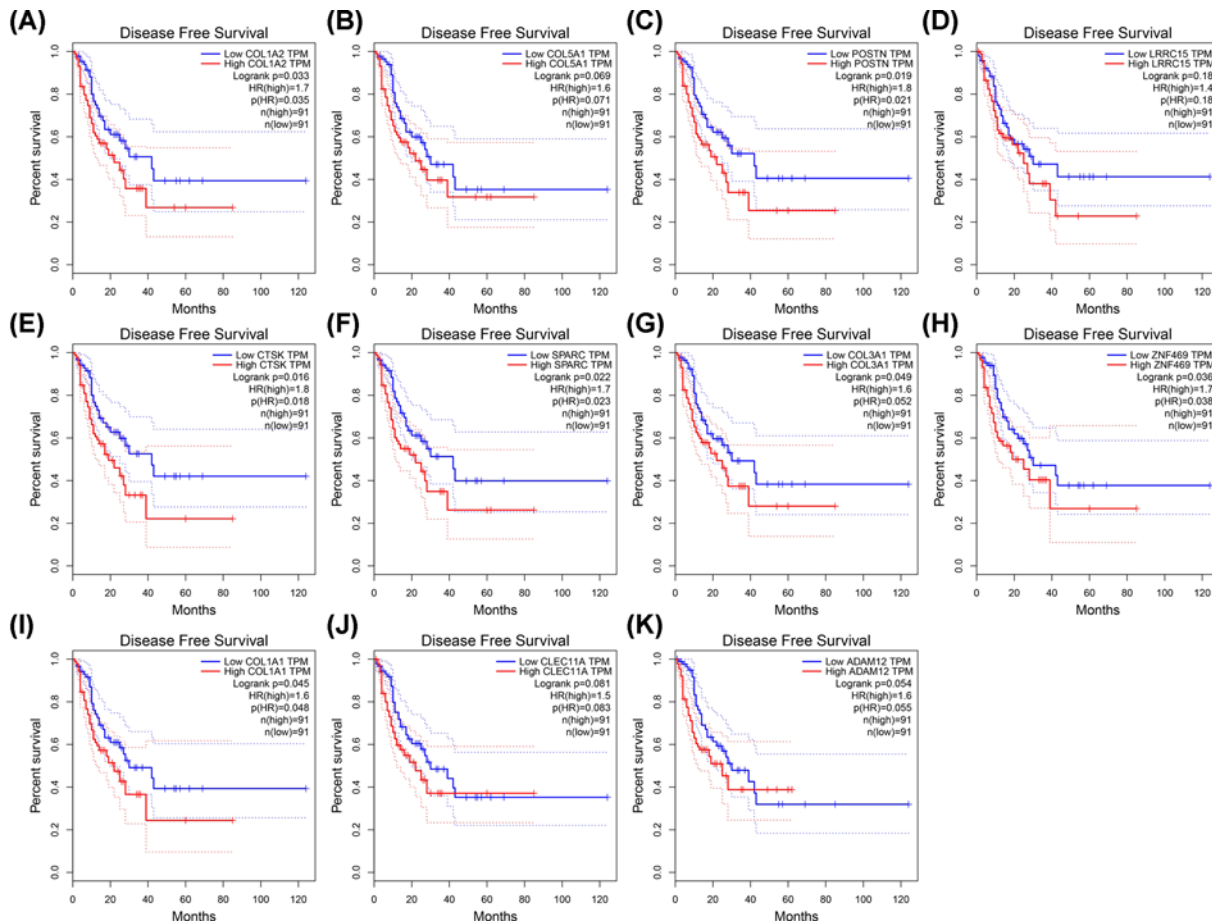


Figure 5. DFS curves of COL1A2 and its similar genes in ESCA from the TCGA database
 (A) COL1A2 (B) COL5A1 (C) POSTN (D) LRRC15 (E) CTSK (F) SPARC (G) COL3A1 (H) ZNF469 (I) COL1A1 (J) CLEC11A (K) ADAM12.
 $P < 0.05$ was considered statistically significant.

instance, Th1 cells, Th2 cells, Tfh cells, Th17 cells, Tregs, and exhausted T cells were also analyzed. Before the correlation analysis, we adjusted it for tumor purity because the tumor purity of clinical samples can impact the immune infiltration analysis. As shown in Table 2, the expression of COL1A2 was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), TAMs (CCL2 and IL10), M1 macrophages (COX2), neutrophils (CD66b, CD11b, and CCR7), DCs (BDCA-4), T-helper (STAT4, GATA3, STAT6, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (CTLA4 and TIM-3). The expression of COL1A1 was significantly correlated with the expression of marker genes from tumor-infiltrating CD8⁺ T cell (CD8B), B cells (CD19 and CD79A), monocytes (CD86 and CD115), TAMs (CCL2 and IL10), M1 macrophages (COX2), neutrophils (CD66b, CD11b, and CCR7), DCs (HLA-DPA1), T-helper (STAT4, STAT5A, STAT6, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (TIM-3). The expression of COL3A1 was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), M1 macrophages (COX2), neutrophils (CD66b and CD11b), DCs (HLA-DQB1 and HLA-DPA1), T-helper (STAT4, GATA3, STAT6, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (CTLA4 and TIM-3). The expression of ZNF469 was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), neutrophils (CD66b and CD11b), T-helper (STAT4, STAT6, STAT5A, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (CTLA4 and TIM-3). The expression of POSTN was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), TAMs (CD68), M1 macrophages (COX2), neutrophils (CD66b, CD11b, and CCR7), T-helper (STAT4, GATA3, STAT6, STAT5A, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β) and exhausted T cells (CTLA4 and TIM-3). All of the results demonstrated that COL1A2 with its similar genes were correlated with the infiltration of the immune cells in ESCA.

Table 2 Correlation analysis between COL1A2 with its similar genes and related genes and markers of immune cells in TIMER

| Description | Gene markers | COL1A2 | | COL1A1 | | COL3A1 | | ZNF469 | | POSTN | |
|-------------------------|--------------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
| | | Purity | | Purity | | Purity | | Purity | | Purity | |
| | | Cor | P | Cor | P | Cor | P | Cor | P | Cor | P |
| CD8 ⁺ T cell | CD8A | -0.070 | 0.353 | -0.127 | 0.090 | -0.055 | 0.466 | -0.031 | 0.681 | -0.044 | 0.561 |
| | CD8B | -0.133 | 0.076 | -0.197 | † | -0.112 | 0.134 | -0.102 | 0.171 | -0.113 | 0.132 |
| T cell (general) | CD3D | -0.063 | 0.399 | -0.116 | 0.122 | -0.045 | 0.544 | -0.050 | 0.507 | -0.068 | 0.363 |
| | CD3E | -0.042 | 0.573 | -0.094 | 0.207 | -0.025 | 0.741 | -0.036 | 0.628 | -0.035 | 0.640 |
| B cell | CD2 | -0.048 | 0.525 | -0.101 | 0.178 | -0.031 | 0.675 | -0.030 | 0.691 | -0.025 | 0.744 |
| | CD19 | -0.156 | * | -0.231 | † | -0.157 | * | -0.155 | * | -0.158 | * |
| Monocyte | CD79A | -0.181 | * | -0.259 | ‡ | -0.153 | * | -0.193 | † | -0.156 | * |
| | CD86 | 0.447 | ‡ | 0.405 | ‡ | 0.455 | ‡ | 0.479 | ‡ | 0.446 | ‡ |
| TAM | CD115 | 0.408 | ‡ | 0.333 | ‡ | 0.431 | ‡ | 0.395 | ‡ | 0.394 | ‡ |
| | CCL2 | 0.482 | ‡ | 0.419 | ‡ | 0.202 | 0.086 | 0.112 | 0.346 | 0.212 | 0.072 |
| M1 Macrophage | CD68 | -0.003 | 0.972 | -0.018 | 0.812 | 0.203 | 0.085 | -0.134 | 0.258 | 0.490 | ‡ |
| | IL10 | 0.387 | ‡ | 0.351 | ‡ | 0.215 | 0.067 | 0.217 | 0.065 | 0.083 | 0.485 |
| M2 Macrophage | INOS | 0.223 | 0.058 | 0.128 | 0.280 | 0.211 | 0.073 | 0.219 | 0.063 | 0.226 | 0.054 |
| | IRF5 | -0.027 | 0.822 | -0.093 | 0.433 | -0.147 | 0.214 | -0.014 | 0.908 | 0.015 | 0.901 |
| Neutrophils | COX2 | 0.382 | ‡ | 0.369 | † | 0.413 | ‡ | 0.147 | 0.216 | 0.355 | † |
| | CD163 | 0.140 | 0.237 | 0.039 | 0.740 | 0.095 | 0.426 | 0.086 | 0.471 | 0.061 | 0.610 |
| NK cell | VSIG4 | 0.163 | 0.168 | 0.093 | 0.434 | 0.090 | 0.449 | 0.064 | 0.591 | 0.055 | 0.645 |
| | MS4A4A | 0.111 | 0.348 | -0.005 | 0.967 | 0.055 | 0.642 | -0.027 | 0.819 | 0.130 | 0.273 |
| DC | CD66b | -0.195 | † | -0.209 | † | -0.185 | * | -0.202 | † | -0.200 | † |
| | CD11b | 0.324 | ‡ | 0.244 | ‡ | 0.335 | ‡ | 0.302 | ‡ | 0.305 | ‡ |
| Th1 | CCR7 | -0.154 | * | -0.241 | † | -0.134 | 0.072 | -0.124 | 0.098 | -0.161 | * |
| | KIR2DL1 | 0.016 | 0.828 | 0.021 | 0.777 | 0.019 | 0.805 | -0.006 | 0.939 | -0.035 | 0.640 |
| Th2 | KIR2DL3 | -0.036 | 0.629 | -0.031 | 0.682 | -0.022 | 0.768 | -0.003 | 0.969 | 0.001 | 0.990 |
| | KIR2DL4 | -0.115 | 0.125 | -0.121 | 0.106 | -0.105 | 0.160 | -0.079 | 0.293 | -0.105 | 0.162 |
| Tfh | KIR3DL1 | -0.010 | 0.898 | -0.023 | 0.754 | -0.022 | 0.769 | 0.048 | 0.524 | -0.067 | 0.372 |
| | KIR3DL2 | 0.006 | 0.938 | 0.017 | 0.817 | 0.002 | 0.978 | 0.005 | 0.949 | -0.005 | 0.946 |
| Th17 | KIR3DL3 | -0.034 | 0.654 | -0.034 | 0.651 | -0.023 | 0.762 | -0.051 | 0.494 | -0.048 | 0.526 |
| | KIR2DS4 | -0.003 | 0.971 | -0.012 | 0.869 | -0.013 | 0.868 | 0.022 | 0.765 | -0.049 | 0.516 |
| Treg | HLA-DPB1 | -0.045 | 0.708 | -0.188 | 0.111 | -0.165 | 0.163 | -0.081 | 0.495 | -0.041 | 0.731 |
| | HLA-DQB1 | -0.152 | 0.199 | -0.222 | 0.059 | -0.254 | * | -0.108 | 0.365 | 0.057 | 0.635 |
| Treg | HLA-DRA | -0.028 | 0.815 | -0.138 | 0.244 | -0.104 | 0.380 | -0.093 | 0.436 | -0.057 | 0.633 |
| | HLA-DPA1 | -0.115 | 0.332 | -0.246 | * | -0.234 | * | -0.161 | 0.173 | 0.012 | 0.918 |
| Treg | BDCA-1 | 0.058 | 0.625 | -0.155 | 0.191 | 0.031 | 0.792 | -0.084 | 0.480 | 0.209 | 0.077 |
| | BDCA-4 | 0.245 | * | 0.185 | 0.117 | 0.186 | 0.116 | 0.134 | 0.260 | 0.058 | 0.626 |
| Treg | CD11c | 0.027 | 0.820 | -0.055 | 0.645 | -0.048 | 0.689 | 0.023 | 0.848 | -0.047 | 0.692 |
| | T-bet | -0.028 | 0.708 | -0.083 | 0.269 | -0.013 | 0.862 | 0.023 | 0.755 | -0.005 | 0.952 |
| Treg | STAT4 | 0.238 | † | 0.172 | * | 0.276 | ‡ | 0.238 | † | 0.292 | ‡ |
| | STAT1 | -0.005 | 0.947 | -0.020 | 0.794 | 0.016 | 0.834 | 0.041 | 0.587 | 0.048 | 0.522 |
| Treg | IFN-γ | 0.017 | 0.817 | -0.001 | 0.990 | 0.024 | 0.745 | 0.061 | 0.415 | 0.008 | 0.918 |
| | TNF-α | 0.041 | 0.581 | 0.063 | 0.400 | 0.036 | 0.636 | 0.112 | 0.133 | 0.018 | 0.811 |
| Treg | GATA3 | 0.154 | * | 0.140 | 0.061 | 0.158 | * | 0.126 | 0.092 | 0.154 | * |
| | STAT6 | -0.311 | ‡ | -0.298 | ‡ | -0.298 | ‡ | -0.270 | ‡ | -0.282 | ‡ |
| Treg | STAT5A | -0.139 | 0.063 | -0.171 | * | -0.114 | 0.127 | -0.160 | * | -0.154 | * |
| | IL13 | 0.091 | 0.227 | 0.047 | 0.530 | 0.109 | 0.143 | 0.076 | 0.309 | 0.055 | 0.462 |
| Treg | BCL6 | 0.288 | ‡ | 0.280 | ‡ | 0.282 | ‡ | 0.344 | ‡ | 0.294 | ‡ |
| | IL21 | 0.019 | 0.802 | 0.002 | 0.977 | 0.022 | 0.768 | 0.064 | 0.394 | 0.066 | 0.381 |
| Treg | STAT3 | -0.126 | 0.091 | -0.143 | 0.055 | -0.121 | 0.105 | -0.129 | 0.085 | -0.103 | 0.171 |
| | IL17A | -0.253 | ‡ | -0.249 | ‡ | -0.225 | † | -0.227 | † | -0.275 | ‡ |
| Treg | FOXP3 | 0.287 | ‡ | 0.200 | † | 0.296 | ‡ | 0.297 | ‡ | 0.286 | ‡ |
| | CCR8 | 0.302 | ‡ | 0.213 | † | 0.323 | ‡ | 0.303 | ‡ | 0.309 | ‡ |
| Treg | STAT5B | 0.107 | 0.152 | 0.032 | 0.669 | 0.123 | 0.101 | 0.123 | 0.101 | 0.105 | 0.160 |
| | TGFβ | 0.395 | ‡ | 0.405 | ‡ | 0.363 | ‡ | 0.449 | ‡ | 0.367 | ‡ |

Continued over

Table 2 Correlation analysis between COL1A2 with its similar genes and related genes and markers of immune cells in TIMER (Continued)

| Description | Gene markers | COL1A2 | | COL1A1 | | COL3A1 | | ZNF469 | | POSTN | |
|-------------------|--------------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
| | | Purity | | Purity | | Purity | | Purity | | Purity | |
| | | Cor | P | Cor | P | Cor | P | Cor | P | Cor | P |
| T cell exhaustion | PD-1 | 0.016 | 0.829 | -0.032 | 0.665 | 0.030 | 0.687 | 0.062 | 0.409 | 0.018 | 0.812 |
| | CTLA4 | 0.184 | * | 0.119 | 0.112 | 0.195 | † | 0.202 | † | 0.189 | * |
| | LAG3 | 0.091 | 0.225 | 0.053 | 0.482 | 0.098 | 0.192 | 0.143 | 0.055 | 0.091 | 0.225 |
| | TIM-3 | 0.426 | ‡ | 0.370 | ‡ | 0.441 | ‡ | 0.445 | ‡ | 0.425 | ‡ |
| | GZMB | 0.005 | 0.952 | -0.009 | 0.900 | -0.002 | 0.976 | 0.070 | 0.351 | -0.014 | 0.850 |

Abbreviations: Cor, R value of Spearman's correlation; Purity, correlation adjusted by purity.

* $P < 0.05$.

† $P < 0.01$.

‡ $P < 0.001$.

Discussion

Over the past decade, there are many studies about bioinformatics analysis that have been published to explore the potential pathological mechanisms in ESCA. However, the mechanism of immune cell infiltration and the genes associated with immune infiltration in ESCA yet remain unclear. We firstly analyzed two subtypes of ESCA respectively, which contained the datasets of ESCC and EAC from GEO to find the common DEGs from the different histologic subtypes. Altogether 128 DEGs including 19 up-regulated genes and 109 down-regulated genes were screened. To demonstrate the functional mechanism of common DEGs in ESCA, we performed GO enrichment and Reactome pathway analysis. The results of GO enrichment and Reactome pathway analysis displayed that the up-regulated DEGs were mainly involved in the regulation of ECM such as extracellular structure organization and ECM organization, while the down-regulated DEGs were mainly involved in the regulation of cornification and keratinocyte differentiation. These results suggest the major variations in ESCA were involved in the remodeling of the tumor microenvironment (TME) and cell differentiation. The TME is an internal and complex environment in the tumor, which includes all kinds of ECM (for instance, collagen and laminin) and stromal cells including fibroblasts, immune cells, and vascular endothelial cells [26,27]. During tumor progression, the TME interacts with tumor cells depending on different kinds of immune cells (for example, DCs, macrophages, and myeloid-derived suppressor cells) to mediate their immunologic tolerance, subsequently impacting the clinical effect of immunotherapy [28,29].

Aiming to find the most significant module involved in immune infiltration, we constructed the PPI network of DEGs and identified the most significant module of up-regulated DEGs by using MCODE. The most significant module of up-regulated DEGs contained COL1A2, COL5A2, COL4A2, LUM, and COL4A1. One of five up-regulated genes, COL1A2, is a subtype of Type I collagen, which can be produced by stromal fibroblasts and cancer cells [30–32]. And COL1A2 plays an important role in the invasion and metastasis of ovarian cancer cells since it promotes the migration of cancer cells by interacting with $\alpha 2\beta 1$ -integrin [33]. However, what roles it plays in immune cell infiltration and immunotherapy about ESCA are still indistinct.

Subsequently, we explored the similar genes of COL1A2 in ESCA by using GEPIA. The top ten similar genes were screened including COL1A1, COL3A1, SPARC, ZNF469, POSTN, CTSK, COL5A1, LRRC15, ADAM12, and CLEC11A. We validated the expressions of COL1A2 and its similar genes in primary ESCA samples compared with normal esophageal squamous epithelium samples, while CTSK, SPARC, and CLEC11A were not significantly up-regulated. Among these genes, the high expression of COL1A2, COL1A1, COL3A1, SPARC, ZNF469, POSTN, and CTSK were negatively correlated with the DFS of ESCA patients. Then these genes (COL1A2, COL1A1, COL3A1, ZNF469, and POSTN) that have been validated and had prognostic value were used for immune-related research. The expression levels of COL1A2, COL1A1, COL3A1, ZNF469, and POSTN were positively correlated with the infiltration levels of macrophages and DCs, and the expression levels of ZNF469 was also positively correlated with the infiltration levels of CD4⁺ T cells. These results indicated these genes might be the candidate genes for assessing the immune infiltration levels in ESCA.

Finally, to further explore the relationship between COL1A2 with its similar genes and the infiltration levels of various immune cells comprehensively, we analyzed the correlation between these genes and immune marker sets. The expression of COL1A2 was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), TAMs (CCL2 and IL10), M1 macrophages (COX2),

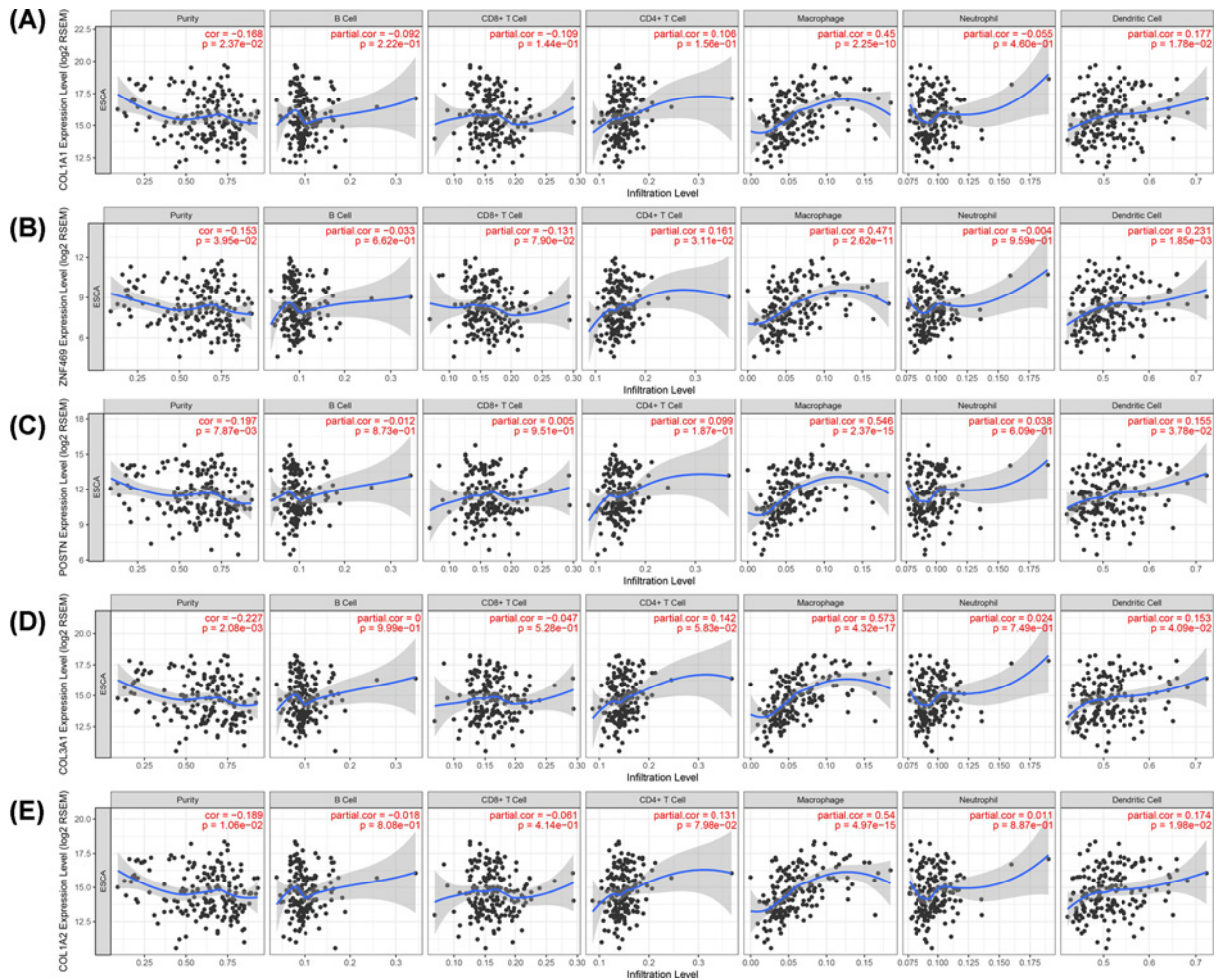


Figure 6. Correlation between the expression levels of COL1A2 with its similar genes and immune infiltration level in ESCA (A) COL1A1 (B) ZNF469 (C) POSTN (D) COL3A1 (E) COL1A2. $P < 0.05$ was considered statistically significant.

neutrophils (CD66b, CD11b, and CCR7), DCs (BDCA-4), T-helper (STAT4, GATA3, STAT6, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (CTLA4 and TIM-3). The expression of COL1A1 was significantly correlated with the expression of marker genes from tumor-infiltrating CD8⁺ T cell (CD8B), B cells (CD19 and CD79A), monocytes (CD86 and CD115), TAMs (CCL2 and IL10), M1 macrophages (COX2), neutrophils (CD66b, CD11b, and CCR7), DCs (HLA-DPA1), T-helper (STAT4, STAT5A, STAT6, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (TIM-3). Type I collagen (COL1A1 and COL1A2) was correlated with the resembling of TAMs in non-small cell lung carcinoma (NSCLC) while it was not correlated with the paucity of T-cell accumulation in pancreatic ductal adenocarcinoma (PDAC), which were consistent with our results [34,35]. The expression of COL3A1 was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), M1 macrophages (COX2), neutrophils (CD66b and CD11b), DCs (HLA-DQB1 and HLA-DPA1), T-helper (STAT4, GATA3, STAT6, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (CTLA4 and TIM-3). COL3A1, a subtype of Type III collagen, which is similar to Type I collagen, and primarily refers to the growth and metastasis of tumors such as osteosarcoma and nasopharyngeal tumors [36,37]. While our results suggested that COL3A1 also can be a potential biomarker for assessing the infiltration of the immune cells in ESCA. The high expression of COL3A1 may promote the growth and metastasis of ESCA by recruiting various immune cells. The expression of ZNF469 was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), neutrophils (CD66b and CD11b), T-helper (STAT4, STAT6, STAT5A, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (CTLA4 and TIM-3). Zinc finger protein (ZNF469) gene mainly refers to the pathological process of cornea-related disease and it is similar to three types of collagen including COL1A1, COL1A2, and COL4A1 [38–40].

Interestingly, we firstly found the important roles of ZNF469 in the assessment of prognosis and immune cell infiltration in ESCA. This gene may be an important biomarker associated with immune infiltration in the ESCA TME. The expression of POSTN was significantly correlated with the expression of marker genes from tumor-infiltrating B cells (CD19 and CD79A), monocytes (CD86 and CD115), TAMs (CD68), M1 macrophages (COX2), neutrophils (CD66b, CD11b, and CCR7), T-helper (STAT4, GATA3, STAT6, STAT5A, BCL6, and IL17A), Treg (FOXP3, CCR8, and TGF β), and exhausted T cells (CTLA4 and TIM-3). POSTN is a soluble ECM protein and it is up-regulated in many cancers such as ESCA, head and neck squamous cell carcinoma, and non-small cell lung cancer [41–43]. Recent studies demonstrated that it was an important biomarker associated with the pathogenesis and prognosis of ESCA, and our results further underlined the important role of POSTN in immune cells infiltration [44,45].

Different from previous studies, we analyzed the common DEGs between the ESCC and EAC to find the most important variations in ESCA and identified the important roles of COL1A2 with its similar genes in influencing the immune infiltration in ESCA. However, some limitations still exist in the present study. Firstly, some technical defects of gene chip and sequencing technology existed here. Secondly, we only analyzed the potential biomarker in ESCA and the specific mechanism of immune cell infiltration induced by these genes are needed to be demonstrated in vivo and in vitro experiments.

In summary, we identified COL1A2 and its similar genes as the potential biomarkers for assessing the infiltration levels of immune cells using a comprehensive analysis of bioinformatics, which could help us better understand the mechanism of tumor immune cell infiltration in ESCA. However, more studies are needed to verify these results.

Data Availability

The gene expression data are available in public databases including GEO (<https://www.ncbi.nlm.nih.gov/geo/>) and TCGA (<https://portal.gdc.cancer.gov>).

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China [grant number 81672688]; the Natural Science Foundation of Hunan Province, China [grant number 2019JJ50359]; the College Students' Innovative Entrepreneurial Training Plan Program of China [grant number 2020105330254]; and the Graduate research innovation project of Central South University [grant number 2020zzts887].

Author Contribution

Z.W., M.C. and Y.Q. designed the present study. Z.W., M.C., Y.Q., Y.Y., Y.H. and X.L. collected the raw data and prepared the analysis. Z.W., M.C. and W.Z. were involved in writing and revising the paper. All authors read and approved the final manuscript.

Abbreviations

BP, biological process; CC, cell component; COL1A2, collagen type I alpha 2 chain; DC, dendritic cell; DEG, differentially expressed gene; DFS, disease-free survival; EAC, esophageal adenocarcinoma; ECM, extracellular matrix; ESCA, esophageal carcinoma; ESCC, esophageal squamous cell carcinoma; FC, fold change; FDR, false discovery rate; GEO, Gene Expression Omnibus; GEPIA, Gene Expression Profiling Interactive Analysis; GO, gene ontology; GTEX, Genotype-Tissue Expression; lncRNA, long noncoding RNA; MF, molecular function; miRNA, microRNA; mRNA, messenger RNA; NK, natural killer; POSTN, Periostin; PPI, protein–protein interaction; TAM, tumor-associated macrophage; TCGA, The Cancer Genome Atlas; Tfh, follicular helper T; Th, T-helper; Th17, T-helper 17; TIMER, Tumor Immune Estimation Resource; TME, tumor microenvironment; Treg, regulatory T cell.

References

- 1 Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. and Jemal, A. (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424, <https://doi.org/10.3322/caac.21492>
- 2 Lagergren, J., Smyth, E., Cunningham, D. and Lagergren, P. (2017) Oesophageal cancer. *Lancet* **390**, 2383–2396, [https://doi.org/10.1016/S0140-6736\(17\)31462-9](https://doi.org/10.1016/S0140-6736(17)31462-9)
- 3 Smyth, E.C., Lagergren, J., Fitzgerald, R.C., Lordick, F., Shah, M.A., Lagergren, P. et al. (2017) Oesophageal cancer. *Nat. Rev. Dis. Primers* **3**, 17048, <https://doi.org/10.1038/nrdp.2017.48>
- 4 Burugu, S., Dancsok, A.R. and Nielsen, T.O. (2018) Emerging targets in cancer immunotherapy. *Semin. Cancer Biol.* **52**, 39–52, <https://doi.org/10.1016/j.semcancer.2017.10.001>

- 5 Sarwat, M. and Yamdagni, M.M. (2016) DNA barcoding, microarrays and next generation sequencing: Recent tools for genetic diversity estimation and authentication of medicinal plants. *Crit. Rev. Biotechnol.* **36**, 191–203, <https://doi.org/10.3109/07388551.2014.947563>
- 6 Rai, G., Rai, R., Saedian, A.H. and Rai, M. (2016) Microarray to deep sequencing: Transcriptome and miRNA profiling to elucidate molecular pathways in systemic lupus erythematosus. *Immunol. Res.* **64**, 14–24, <https://doi.org/10.1007/s12026-015-8672-y>
- 7 Mantione, K.J., Kream, R.M., Kuzelova, H., Ptacek, R., Raboch, J., Samuel, J.M. et al. (2014) Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. *Med. Sci. Monit. Basic Res.* **20**, 138–142, <https://doi.org/10.12659/MSMBR.892101>
- 8 Yue, Y., Song, M., Qiao, Y., Li, P., Yuan, Y., Lian, J. et al. (2017) Gene function analysis and underlying mechanism of esophagus cancer based on microarray gene expression profiling. *Oncotarget* **8**, 105222–105237, <https://doi.org/10.18632/oncotarget.22160>
- 9 Liu, H., Zhang, Q., Lou, Q., Zhang, X., Cui, Y., Wang, P. et al. (2019) Differential analysis of lncRNA, miRNA and mRNA expression profiles and the prognostic value of lncRNA in esophageal cancer. *Pathol. Oncol. Res.* **26**, 1029–1039, <https://doi.org/10.1007/s12253-019-00655-8>
- 10 Zeng, J.H., Xiong, D.D., Pang, Y.Y., Zhang, Y., Tang, R.X., Luo, D.Z. et al. (2017) Identification of molecular targets for esophageal carcinoma diagnosis using miRNA-seq and RNA-seq data from the Cancer Genome Atlas: a study of 187 cases. *Oncotarget* **8**, 35681–35699, <https://doi.org/10.18632/oncotarget.16051>
- 11 Gautier, L., Cope, L., Bolstad, B.M. and Irizarry, R.A. (2004) Affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315
- 12 Smyth, G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, e3
- 13 Mi, H., Muruganujan, A., Ebert, D., Huang, X. and Thomas, P.D. (2019) PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **47**, D419–D426
- 14 Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J. et al. (2019) STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613
- 15 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D. et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504
- 16 Bandettini, W.P., Kellman, P., Mancini, C., Booker, O.J., Vasu, S., Leung, S.W. et al. (2012) MultiContrast Delayed Enhancement (MCOE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. *J. Cardiovasc. Magn. Reson.* **14**, 83
- 17 Chandrashekar, D.S., Bashel, B., Balasubramanya, S., Creighton, C.J., Ponce-Rodriguez, I., Chakravarthi, B. et al. (2017) UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* **19**, 649–658
- 18 Tang, Z., Li, C., Kang, B., Gao, G., Li, C. and Zhang, Z. (2017) GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**, W98–W102
- 19 Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J.S. et al. (2017) TIMER: A web server for comprehensive analysis of Tumor-Infiltrating immune cells. *Cancer Res.* **77**, e108–e110
- 20 Danaher, P., Warren, S., Dennis, L., D'Amico, L., White, A., Disis, M.L. et al. (2017) Gene expression markers of Tumor Infiltrating Leukocytes. *J. Immunother. Cancer* **5**, 18
- 21 Siemers, N.O., Holloway, J.L., Chang, H., Chasalow, S.D., Ross-Macdonald, P.B., Voliva, C.F. et al. (2017) Genome-wide association analysis identifies genetic correlates of immune infiltrates in solid tumors. *PLoS ONE* **12**, e179726, <https://doi.org/10.1371/journal.pone.0179726>
- 22 Sousa, S. and Maatta, J. (2016) The role of tumour-associated macrophages in bone metastasis. *J. Bone Oncol.* **5**, 135–138, <https://doi.org/10.1016/j.jbo.2016.03.004>
- 23 Fang, S., Dai, Y., Mei, Y., Yang, M., Hu, L., Yang, H. et al. (2019) Clinical significance and biological role of cancer-derived Type I collagen in lung and esophageal cancers. *Thorac. Cancer* **10**, 277–288, <https://doi.org/10.1111/1759-7714.12947>
- 24 Miyashita, T., Tajima, H., Shah, F.A., Oshima, M., Makino, I., Nakagawara, H. et al. (2014) Impact of inflammation-metaplasia-adenocarcinoma sequence and inflammatory microenvironment in esophageal carcinogenesis using surgical rat models. *Ann. Surg. Oncol.* **21**, 2012–2019, <https://doi.org/10.1245/s10434-014-3537-5>
- 25 Yoshihara, K., Shahmoradgol, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W. et al. (2013) Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612, <https://doi.org/10.1038/ncomms3612>
- 26 Wu, T. and Dai, Y. (2017) Tumor microenvironment and therapeutic response. *Cancer Lett.* **387**, 61–68, <https://doi.org/10.1016/j.canlet.2016.01.043>
- 27 Karlou, M., Tzelepi, V. and Efstathiou, E. (2010) Therapeutic targeting of the prostate cancer microenvironment. *Nat. Rev. Urol.* **7**, 494–509, <https://doi.org/10.1038/nrurol.2010.134>
- 28 Gao, S., Yang, D., Fang, Y., Lin, X., Jin, X., Wang, Q. et al. (2019) Engineering nanoparticles for targeted remodeling of the tumor microenvironment to improve cancer immunotherapy. *Theranostics* **9**, 126–151, <https://doi.org/10.7150/thno.29431>
- 29 Zou, W. (2005) Immunosuppressive networks in the tumour environment and their therapeutic relevance. *Nat. Rev. Cancer* **5**, 263–274, <https://doi.org/10.1038/nrc1586>
- 30 Steinberg, J. (1973) The turnover of collagen in fibroblast cultures. *J. Cell Sci.* **12**, 217–234
- 31 Steinberg, J. (1978) Collagen turnover and the growth state in 3T6 fibroblast cultures. *Lab. Invest.* **39**, 491–496
- 32 Kauppila, S., Stenback, F., Risteli, J., Jukkola, A. and Risteli, L. (1998) Aberrant type I and type III collagen gene expression in human breast cancer in vivo. *J. Pathol.* **186**, 262–268, [https://doi.org/10.1002/\(SICI\)1096-9896\(1998110\)186:3%3c262::AID-PATH191%3e3.0.CO;2-3](https://doi.org/10.1002/(SICI)1096-9896(1998110)186:3%3c262::AID-PATH191%3e3.0.CO;2-3)
- 33 Fishman, D.A., Kearns, A., Chilukuri, K., Bafetti, L.M., O'Toole, E.A., Georgacopoulos, J. et al. (1998) Metastatic dissemination of human ovarian epithelial carcinoma is promoted by alpha2beta1-integrin-mediated interaction with type I collagen. *Invasion Metastasis* **18**, 15–26, <https://doi.org/10.1159/000024495>
- 34 Rebelo, S.P., Pinto, C., Martins, T.R., Harrer, N., Estrada, M.F., Loza-Alvarez, P. et al. (2018) 3D-3-culture: A tool to unveil macrophage plasticity in the tumour microenvironment. *Biomaterials* **163**, 185–197, <https://doi.org/10.1016/j.biomaterials.2018.02.030>

- 35 Carstens, J.L., Correa, D.S.P., Yang, D., Barua, S., Wang, H., Rao, A. et al. (2017) Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer. *Nat. Commun.* **8**, 15095, <https://doi.org/10.1038/ncomms15095>
- 36 Xu, W., Li, Z., Zhu, X., Xu, R. and Xu, Y. (2018) MiR-29 family inhibits resistance to methotrexate and promotes cell apoptosis by targeting COL3A1 and MCL1 in osteosarcoma. *Med. Sci. Monit.* **24**, 8812–8821, <https://doi.org/10.12659/MSM.911972>
- 37 Qiu, F., Sun, R., Deng, N., Guo, T., Cao, Y., Yu, Y. et al. (2015) MiR-29a/b enhances cell migration and invasion in nasopharyngeal carcinoma progression by regulating SPARC and COL3A1 gene expression. *PLoS ONE* **10**, e120969, <https://doi.org/10.1371/journal.pone.0120969>
- 38 Yu, X., Chen, B., Zhang, X. and Shentu, X. (2017) Identification of seven novel ZNF469 mutations in keratoconus patients in a Han Chinese population. *Mol. Vis.* **23**, 296–305
- 39 Micheal, S., Siddiqui, S.N., Zafar, S.N., Gabriela, N.I., Khan, M.I. and Bergen, A. (2019) Identification of a novel ZNF469 mutation in a pakistani family with brittle cornea syndrome. *Cornea* **38**, 718–722, <https://doi.org/10.1097/ICO.0000000000001828>
- 40 Abu, A., Frydman, M., Marek, D., Pras, E., Nir, U., Reznik-Wolf, H. et al. (2008) Deleterious mutations in the Zinc-Finger 469 gene cause brittle cornea syndrome. *Am. J. Hum. Genet.* **82**, 1217–1222, <https://doi.org/10.1016/j.ajhg.2008.04.001>
- 41 Michaylira, C.Z., Wong, G.S., Miller, C.G., Gutierrez, C.M., Nakagawa, H., Hammond, R. et al. (2010) Periostin, a cell adhesion molecule, facilitates invasion in the tumor microenvironment and annotates a novel tumor-invasive signature in esophageal cancer. *Cancer Res.* **70**, 5281–5292, <https://doi.org/10.1158/0008-5472.CAN-10-0704>
- 42 Sasaki, H., Lo, K.M., Chen, L.B., Auclair, D., Nakashima, Y., Moriyama, S. et al. (2001) Expression of Periostin, homologous with an insect cell adhesion molecule, as a prognostic marker in non-small cell lung cancers. *Jpn. J. Cancer Res.* **92**, 869–873, <https://doi.org/10.1111/j.1349-7006.2001.tb01174.x>
- 43 Kudo, Y., Ogawa, I., Kitajima, S., Kitagawa, M., Kawai, H., Gaffney, P.M. et al. (2006) Periostin promotes invasion and anchorage-independent growth in the metastatic process of head and neck cancer. *Cancer Res.* **66**, 6928–6935, <https://doi.org/10.1158/0008-5472.CAN-05-4540>
- 44 Zhang, H., Zhong, J., Tu, Y., Liu, B., Chen, Z., Luo, Y. et al. (2019) Integrated bioinformatics analysis identifies hub genes associated with the pathogenesis and prognosis of esophageal squamous cell carcinoma. *Biomed Res. Int.* **2019**, 2615921, <https://doi.org/10.1155/2019/2615921>
- 45 Heidari, P., Esfahani, S.A., Turker, N.S., Wong, G., Wang, T.C., Rustgi, A.K. et al. (2015) Imaging of secreted extracellular periostin, an important marker of invasion in the tumor microenvironment in esophageal cancer. *J. Nucl. Med.* **56**, 1246–1251, <https://doi.org/10.2967/jnumed.115.156216>