**Name of Journal:** World Journal of Gastrointestinal Oncology  
**Manuscript NO:** 90008  
**Manuscript Type:** ORIGINAL ARTICLE

**Basic Study**
Transcriptome sequencing reveals novel biomarkers and immune cell infiltration in esophageal tumorigenesis

Sun JR *et al.*. Novel biomarkers for esophageal tumorigenesis
Abstract

BACKGROUND
Esophageal squamous cell carcinoma (ESCC) is one of the most common malignancies worldwide, and its development comprises a multistep process from intraepithelial neoplasia (IN) to carcinoma (CA). However, the critical regulators and underlying molecular mechanisms remain largely unknown.

AIM
To explore the genes and infiltrating immune cells in the microenvironment that are associated with the multistage progression of ESCC to facilitate diagnosis and early intervention.

METHODS
A mouse model mimicking the multistage development of ESCC was established by providing water containing 4-nitroquinoline 1-oxide (4NQO) to C57BL/6 mice. Moreover, we established a control group without 4NQO treatment of mice. Then, transcriptome sequencing was performed for esophageal tissues from patients with different pathological statuses, including low-grade IN (LGIN), high-grade IN (HGIN), and CA, and controlled normal tissue (NOR) samples. Differentially expressed genes (DEGs) were identified in the LGIN, HGIN, and CA groups, and the biological functions of the DEGs were analyzed via Gene Ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses. The CIBERSORT algorithm was used to detect the pattern of immune cell infiltration. Immunohistochemistry (IHC) was also conducted to validate our results. Finally, the Luminex multiplex cytokine analysis was utilized to measure the serum cytokine levels in the mice.

RESULTS
Compared with those in the NOR group, a total of 681541, and 840 DEGs were obtained in the LGIN, HGIN, and CA groups, respectively. Using the intersection of the three
sets of DEGs, we identified 86 genes as key genes involved in the development of ESCC. Enrichment analysis revealed that these genes were enriched mainly in the keratinization, epidermal cell differentiation, and interleukin (IL)-17 signaling pathways. CIBERSORT analysis revealed that, compared with those in the NOR group, M0 and M1 macrophages in the 4NQO group showed stronger infiltration, which was validated by IHC. Serum cytokine analysis revealed that, compared with those in the NOR group, IL-1β and IL-6 were upregulated, while IL-10 was downregulated in the LGIN, HGIN, and CA groups. Moreover, the expression of the representative key genes, such as S100a8 and Krt6b, was verified in external human samples, and the results of immunohistochemical staining were consistent with the findings in mice.

CONCLUSION

We identified a set of key genes represented by S100a8 and Krt6b and investigated their potential biological functions. In addition, we found that macrophage infiltration and abnormal alterations in the levels of inflammation-associated cytokines, such as IL-1β, IL-6, and IL-10, in the peripheral blood may be closely associated with the development of ESCC.

Key Words: Esophageal squamous cell carcinoma; Intraepithelial neoplasia; Tumorigenesis; Transcriptome sequencing; Biomarkers; Immune cell infiltration; 4-nitroquinoline 1-oxid

Sun JR, Chen DM, Huang R, Wang RT, Jia LQ. Transcriptome sequencing reveals novel biomarkers and immune cell infiltration in esophageal tumorigenesis. World J Gastrointest Oncol 2024; In press

Core Tip: The development of esophageal squamous cell carcinoma (ESCC) involves a stepwise progression from intraepithelial neoplasia (IN) to carcinoma (CA). Examining the genes and immune cell infiltration in the microenvironment associated with the
multi-stage progression of ESCC is essential for facilitating diagnosis and early intervention. A total of 86 crucial genes linked with ESCC development were discovered using transcriptome sequencing. These genes were enriched in pathways related to keratinization, epidermal cell differentiation, and interleukin (IL)-17 signaling. Additionally, the infiltration of macrophages and abnormal alterations of inflammation-associated cytokines (e.g., IL-1β, IL-6, and IL-10) in peripheral blood may be strongly linked with the development of ESCC.

INTRODUCTION

Esophageal cancer (EC) is the seventh most prevalent cancer worldwide. In 2020, there were 604100 new cases and 544000 deaths due to EC[1]. EC can be histopathologically classified into two groups: Squamous cell carcinoma (SCC) and adenocarcinoma (AC). SCC is the most common form of EC and accounts for 80% of all cases globally[2]. SCC originates from low-grade intraepithelial neoplasia (LGIN) and progresses to high-grade IN (HGIN) and carcinoma (CA). Currently, the mechanisms underlying the development and progression of ESCC are largely unknown. Therefore, studying the mechanisms underlying ESCC development and identifying biomarkers are crucial for timely detection, diagnosis, and treatment.

Individual differences may cause heterogeneity in esophageal samples collected during cross-sectional studies, potentially leading to biased results. However, this issue can be addressed by using animal models. The chemical carcinogen 4-nitroquinoline 1-oxide (4NQO) can induce the development of ESCC in mice, similar to the development of ESCC in humans[3,4]. This multistage esophageal carcinogenesis may lead to the discovery of the molecular traits of ESCC through RNA sequencing (RNA-seq) of the transcriptome.

A multistage mouse model of ESCC tumorigenesis was established, and RNA-seq was performed in the present study. Transcriptomic profiling revealed the molecular features of esophageal tumor formation, illustrated the role of these features in determining pathological grade, and revealed a set of key genes associated with the
development of ESCC. Furthermore, we studied the pattern of immune cell infiltration in ESCC. These findings illuminate the molecular changes that occur during the development of ESCC in animal models and may have implications for managing human ESCC.

MATERIALS AND METHODS

Materials

4NQO was acquired from Sigma-Aldrich. Antibodies against S100a8, Krt6b, Ki-67, and CD68 were purchased from Cell Signaling Technology. Antibodies against CD80 and CD86 were obtained from Proteintech. Antibodies against CK5 and p40 were purchased from Servicebio.

Induction of esophageal tumorigenesis model

A total of 35 mice were included in this study, 30 of which were assigned to the 4NQO treatment group and 5 to the control group. Animals of eight-week-old C57BL6 mice were purchased from the Beijing Changyang Xishan Breeding Center and were maintained in a local housing facility under a controlled conditions (23 ± 1 °C, 50% ± 10% humidity and 12-12 h light-dark cycle). Mice were administered 4NQO (100 µg/mL) in their drinking water for 16 wk to stimulate esophageal carcinogenesis. Freshly-prepared drinking water containing the carcinogen was administered once a week, and the mice had free access to drinking water during treatment. After 16 wk of carcinogen exposure, the mice received sterile water instead of 4NQO-containing drinking water until the end of the 24th wk. In addition, in the control group in which mice were not treated with 4NQO and were fed until the end of the 24th wk.

The animals were euthanized according to the accepted method of euthanasia as defined by the American Veterinary Medical Association Guidelines on Euthanasia-Approved Euthanasia Method, 2013. At the end of the 24th wk, the animals were euthanized with ketamine (100 mg/kg body weight) intraperitoneally as well as via a secondary method of cervical dislocation.
Immediately after euthanasia, the esophagus was isolated and rinsed with PBS. The esophagus was frozen at -80 °C for RNA-seq analysis and subsequently fixed in 10% formalin. Paraffin-embedded tissues were sectioned to a thickness of 4 μm for hematoxylin and eosin (HE) staining and immunohistochemistry (IHC).

Two pathologists identified the esophageal lesions based on previously established histopathological criteria[5]. In summary, the normal (NOR) stage consists of a well-oriented, stratified epithelium with a basal zone and a superficial zone. IN is characterized by loss of polarity in epithelial cells, hyperchromatic nuclei with pleomorphism, and increased or abnormal mitotic activity. In LGIN, these abnormalities occur in less than 50% of the epithelium. In high-grade IN (HGIN), however, these changes are present in 50% or more of the epithelium. Lesions that exhibit abnormal changes throughout the entire thickness of the epithelium or invade subepithelial tissues are classified as CA[6].

After clarifying the esophageal pathology via HE staining, we randomly selected three samples each of esophageal tissues portraying the LGIN, HGIN, and CA pathological stages in the 4NQO treatment group and three samples of NOR in the control group for subsequent transcriptome sequencing.

**RNA-seq**

Total RNA was extracted from the tissue using TRIzol®Reagent according to the manufacturer’s instructions. RNA quality was assessed using a 5300 Bioanalyzer (Agilent) and quantified using an ND-2000 system (NanoDrop Technologies). The sequencing library was constructed only with only high-quality RNA samples (OD260/280 = 1.8-2.2, OD260/230 ≥ 2.0, RIN ≥ 6.5, 28S:18S ≥ 1.0, > 1 μg).

RNA purification, reverse transcription, library construction, and sequencing were performed at Shanghai Majorbio Biopharm Biotechnology Co., Ltd., following standard protocols specified by the manufacturer (Illumina, San Diego, CA). An RNA-seq transcriptome library was prepared using 1 μg of total RNA following the Illumina® Stranded mRNA Prep, Ligation Kit from Illumina (San Diego, CA). Briefly, messenger
RNA was isolated using the polyA selection method with oligo (dT) beads. Then, the samples were fragmented using a fragmentation buffer. Next, double-stranded complementary DNA (cDNA) was synthesized using a SuperScript double-stranded cDNA synthesis kit (Invitrogen, CA) and random hexamer primers (Illumina). The synthesized cDNA was then subjected to end-repair, phosphorylation, and ‘A’ base addition following Illumina’s library construction protocol. Libraries were selected based on the size of the cDNA target fragments of (300 bp) on a 2% low range ultra agarose. The products were subsequently amplified using Phusion DNA polymerase (NEB) with 15 polymerase chain reaction (PCR) cycles. After being quantified with a Qubit 4.0, the paired-end RNA-seq sequencing library was sequenced using a NovaSeq 6000 sequencer with a $2 \times 150$ bp read length.

The raw paired-end reads were trimmed and quality controlled by fastp, with default parameters. Then, the clean reads were separately aligned to the reference genome in orientation mode with HISAT2 software\textsuperscript{[7,8]}. The mapped reads of each sample were assembled by StringTie \textit{via} a reference-based approach to estimate the expression level of each gene in the sample\textsuperscript{[9]}. Ultimately, StringTie combines the expression estimates of the individual transcripts and outputs the expression level of each gene, \textit{i.e.}, the reads per million transcripts (TPM). The resulting TPM values were subsequently used for the subsequent differential gene expression analysis.

\textit{Differential expression analysis and functional enrichment}

For identification of differentially expressed genes (DEGs) between different samples, the expression level of each transcript was calculated according to the TPM method. Transcript levels were determined using the TPM method. Gene abundances were quantified using RSEM\textsuperscript{[10]}. Differential expression analysis was performed using the ‘limma’ package in the R software\textsuperscript{[11]}. DEGs with a $|\log_2 \text{fold change}| > 1$ and an adjusted $P$-value $< 0.05$ were considered significant DEGs. In addition, functional-enrichment analyses, including Gene Ontology (GO) and Kyoto Genome Encyclopedia (KEGG) analyses, were performed with the ‘clusterProfiler’ package in the R
software\textsuperscript{[12]}, to identify which DEGs were significantly enriched in GO terms and metabolic pathways compared with the whole-transcriptome background. A Bonferroni-corrected \( P \)-value \( \leq 0.05 \) was considered as the threshold for significant differences.

**Immune cell infiltration**

We used CIBERSORT, a deconvolution algorithm supported by the R package, to calculate the proportions of 22 immune cell types in each sample. The algorithm utilizes gene expression profiles to quantify immune infiltration and the reference signature matrix for CIBERSORT analysis is LM22\textsuperscript{[13]}. First, we acquired the correspondence table for homologous genes of mice and humans from the Ensemble website (https://asia.ensembl.org/index.html). Then, we converted the mouse gene into a homologous human gene. We subsequently conducted a CIBERSORT analysis. Moreover, the results were validated by IHC.

**Relative quantitative real-time PCR**

The expression levels of key genes (S100a8 and Krt6b) were measured in the esophageal tissues of mice (NOR = 3, LGIN = 3, HGIN = 6, CA = 6). Total RNA was extracted from the esophageal samples using TRIzol reagent (Life Technologies, United States) according to the manufacturers’ instructions. The synthesis of cDNA was conducted according to the RNA PCR kit protocol (TaKaRa Bio, China). Then, the cDNAs were subjected to PCR (for S100a8, Krt6b, and GAPDH) with SYBR Green reagents (Thermo Fisher Scientific, United States) under the reaction conditions. The primer sequences are shown in Supplementary Table 1. The relative value of mRNA expression was calculated by the comparative \( \Delta \Delta Ct \) method using GAPDH as a reference gene.

**IHC**

Protein expression in the esophageal tissues was assessed using immunohistochemical staining (NOR = 3, LGIN = 3, HGIN = 6, CA = 6). Following fixation with 10% formalin,
the esophageal tissues were cut into 4 μm thick sections. The sections were dewaxed and rehydrated using xylene and ethanol. Endogenous peroxidase activity was inhibited by treating the sections with 3% H₂O₂ for 15 min. The sections were first blocked with 1% goat serum, and then antigen retrieval was performed by incubating the sections with Tris-EDTA buffer (pH = 9.0) for 10 min. Subsequently, the sections were washed three times for 3 min each and incubated with primary antibodies at 4 °C overnight. Thereafter, the samples were incubated with a biotinylated secondary antibody conjugated with streptavidin-biotinylated horseradish peroxidase.

**Validation in mouse and human ESCC samples**
First, we determined the mRNA and protein expression levels of key genes (S100a8 and Krt6b) in esophageal tissue samples from mice using relative quantitative real-time PCR (qPCR) and IHC, respectively. We then validated these results using formalin-fixed paraffin-embedded sections of esophageal precursor lesions collected from 31 patients at Cixian Cancer Hospital (Hebei Province, China) between 2021 and 2022; the sections were from the NOR (n = 5), LGIN (n = 10), HGIN (n = 10), and CA (n = 6) groups (Supplementary Table 2). The expression levels of S100a8 and Krt6b were determined through IHC staining of human specimens, and all the results were subjected to quantitative analysis using ImageJ. Approval for this study was obtained from the Institutional Ethics Committee of Cixian Cancer Hospital and China-Japan Friendship Hospital. Clinical information was retrieved from patients’ medical records.

**Cytokine assay**
We determined the serum concentrations of interleukin (IL)-1β, IL-2, IL-6, IL-10, IL-12, and tumor necrosis factor-α in mice using a Luminex multiplex cytokine kit following the instructions provided by the manufacturer.

**Statistical analysis**
All the statistical analyses were conducted with R software (version 4.1.0). Continuous data were analyzed by ANOVA (normally distributed data) or the Kruskal-Wallis test (non-normally distributed data). Categorical data were analyzed by the χ² methods. P < 0.05 was regarded as statistically significant.

RESULTS

Development of esophageal tumors
Here, we established a mouse model of EC using 4NQO (Figure 1A). In this study, 30 mice were treated with 4NQO while an additional five mice were not treated and served as the control group. By the end of the 24th wk, in the 4NQO-treated group, five mice had died, and the remaining 25 displayed varying degrees of malignant esophageal transformation. We observed a smooth surface and normal thickness in the NOR group, and the tissue did not show any unusual convexity or concavity. However, in the 4NQO group, the esophageal tissue exhibited thick, rough, white patches and palpable protuberances (Supplementary Figure 1). The esophageal epithelium in the 4NQO group exhibited LGIN (n = 3), HGIN (n = 16), and CA (n = 6, 1 CA in situ, 5 invasive CAs). HE staining revealed that the esophageal mucosal layer was thickened, the nuclei of the epithelial cells were enlarged and hyperchromatic, and the polarity of the cells disappeared in the LGIN and HGIN tissues. Furthermore, the esophageal mucosal layer was thickened in the LGIN and HGIN groups (Figure 1B). In addition, the expression levels of the malignant marker Ki-67 increased with the progression of esophageal pathology (Figure 1C). Pathology markers for SCC, specifically p40 (Figure 1D) and CK5 (Figure 1E), were found to be overexpressed in the LGIN, HGIN, and CA groups. These findings indicate that the induction of esophageal carcinogenesis by 4NQO in mice can replicate the pathological process implicated in the onset of ESCC in humans. Moreover, the tumor histology observed in this model predominantly revealed SCC.

Identifying DEGs during esophageal tumorigenesis
We performed a systemic differential expression analysis. Compared with those in the NOR group, 681 DEGs - 397 upregulated and 284 downregulated genes - were identified in the LGIN group. Similarly, 541 DEGs - 397 upregulated and 144 downregulated genes - were identified in the HGIN group and 840 DEGs - 634 upregulated and 206 downregulated genes - were found in the CA group (Figure 2A-C, Supplementary Tables 3-5).

**Identification of key genes involved in the development of ESCC and functional enrichment analysis**

The intersection between DEGs in the LGIN, HGIN, and CA groups was detected to identify the common key genes involved in the development of ESCC. In particular, to avoid confusion, we initially identified the upregulated DEGs in each group and performed an intersection analysis of these DEGs in the LGIN, HGIN, and CA groups. A total of 58 intersecting genes were identified (Figure 2D). We subsequently performed the same analysis for the downregulated DEGs and identified 28 intersecting genes (Figure 2E). Finally, 86 genes were identified that could be regarded as key genes in the development of ESCC (Supplementary Table 6). The most significant key genes, Krt6b, S100a8, Dfeb3, Sprr2h, Krt16, and Sprr2i, were annotated in the heatmap (Figure 2F).

Furthermore, these 86 genes were analyzed via GO and KEGG analyses to explore their functions. Among the biological process terms, 40 terms were enriched according to the GO enrichment analysis. These genes were enriched primarily in keratinization, keratinocyte differentiation, epidermal cell differentiation, and epidermis development (Figure 3A). Among the cellular component terms, 14 terms were enriched mainly in the cornified envelope, peptidase inhibitor, and apical part of the cell (Figure 3B). With respect to molecular function, which includes 26 terms, these genes were found to be enriched in structural constituents of the skin epidermis, endopeptidase activity, and peroxidase activity (Figure 3C). Here, we present only the 15 most enriched terms. The KEGG analysis revealed that these genes were predominantly involved in pathways
such as IL-17 signaling, primary bile acid biosynthesis, and nitrogen metabolism (Figure 3D).

Infiltration of immune cells in the ESCC microenvironment
CIBERSORT software was used to estimate the abundance of immune cells in each sample[13]. Analysis of the abundance of immune cells in the different groups revealed that compared to those in the normal group, M0 macrophages ($P < 0.05$) and M1 macrophages ($P < 0.05$) were more abundant in the 4NQO group (Figure 4A). The expression of M0 (CD68) and M1 (CD80 and CD86) markers was validated via immunohistochemical staining. The results showed that compared with the normal group, the 4NQO group with different pathological statuses (LGIN, HGIN, and CA) had a greater abundance of M0 and M1 macrophages than the normal group (Figure 4B-D).

Expression of key genes in mouse and human samples
The top 2 key genes, Krt6b and S100a8, were selected for validation by qPCR and IHC. The qPCR results showed that both genes were highly expressed in the LGIN, HGIN, and CA groups compared with the NOR group (Figure 5A and B), consistent with the results of transcriptomic sequencing (Figure 5C and D). The protein levels of S100a8 and Krt6b were overexpressed in the LGIN, HGIN, and CA groups as shown by the IHC results for mouse samples (Figure 6A and B). Furthermore, external human samples were utilized to confirm our findings in mice, and the IHC results indicated good consistency (Figure 6C and D). All the results were consistent, and the results are shown in Figure 6E-H.

The serum cytokine levels in mice
The serum cytokine analysis indicated that the serum levels of IL-1β and IL-6 were increased and the serum levels of IL-10 were decreased in the LGIN, HGIN, and CA
groups compared to the NOR group (Figure 7A-C). However, no significant differences in the expression of other cytokines were observed (Figure 7D-F).

**DISCUSSION**

ESCC is a multifactorial disease resulting from a complex interplay of genetic and environmental factors that leads to the gradual malignant transformation of esophageal epithelial cells. The progression of ESCC is a lengthy process, and precancerous lesions may persist for a prolonged period, making it difficult to obtain human samples representing different stages of carcinogenesis. To mimic the human ESCC, researchers have generated murine ESCC models _via_ the exposure to 4NQO, a potent carcinogenic chemical[^3]. The findings of our mouse model indicate its effectiveness in precisely pinpointing the genes implicated in the transition of normal epithelial cells to malignant epithelial cells.

In the present study, our results indicated that 4NQO can successfully induce atypical hyperplasia or cancerous changes in esophageal tissues at 24 wk. The multistage disease process in the murine model was similar to that in humans. The proliferative capacity of esophageal basal cells increases with disease progression, indicating a malignant transformation. Through transcriptome sequencing analysis, we identified 397 key genes that play crucial roles in the development of ESCC. Subsequently, functional enrichment analysis was performed to explore the biological function of these key genes. The results demonstrated that these genes were related to keratinization, keratinocyte differentiation, epidermal cell differentiation, and epithelial cell proliferation. Keratin constitutes the intermediate filament in epithelial cells. In cancer, keratin is extensively used as a diagnostic tumor marker, as epithelial malignancies largely express specific keratins associated with their origin. Several studies have shown that keratin may participate in cancer cell invasion and metastasis by regulating epithelial tumorigenesis[^4]. The epidermis is formed and maintained by intricate genetic networks that link cellular differentiation processes, enzymatic activities, and cellular junctions. Disruption of these networks affects the balance between keratinocyte
proliferation and differentiation, impairing epithelial integrity and leading to epidermal hyperproliferation and SCC. Recent studies on wound-induced and inflammation-mediated cancers in mice have identified the dysregulation of core barrier components as a major driver of tumor growth. Thus, it was hypothesized that the loss of barrier-related genes induces epithelial barrier dysfunction, inflammation-mediated epidermal hyperplasia, and carcinogenesis over time. This emerging perspective suggests that localized barrier impairment can be considered a hallmark of primary lesions in epidermal SCC under specific genetic circumstances\(^{15}\).

Two representative key genes, namely Krt6b and S100a8, exhibited the most significant differences in expression and were validated by IHC. These genes are considered to play important roles in esophageal carcinogenesis. Krt6b, a member of the keratin gene family, can be detected in bladder cancer-derived exosomes and plays an important role in epithelial-mesenchymal transition and immune responses in bladder cancer\(^{16}\). Another study showed that Krt6b is upregulated in metastatic bladder tumors and that its expression in bladder cancer samples is positively correlated with tumor grade. In addition, Krt6b potentially functions as a marker for diagnosing and predicting the progression of melanoma\(^{17}\). S100a8 is a Ca\(^{2+}\)-binding protein that belongs to the S100 family. During inflammation, S100A8/A9 is actively released and modulates the inflammatory response by recruiting leukocytes and inducing cytokine secretion\(^{18}\). Numerous cancers originate from sites of infection, persistent irritation, and inflammation. The inflammatory microenvironment plays an essential role in the overproduction of inflammatory factors, thereby inducing neoplastic processes\(^{19}\). S100a8 also participates in several signaling pathways in tumor cells and contributes to tumor development, growth, and metastasis by interfering with tumor metabolism and regulating the tumor microenvironment\(^{20}\). For instance, coculturing ESCC cells with macrophages enhances S100A8/9 expression and release by ESCC cells, promoting ESCC progression through the Akt and p38 MAPK signaling pathways\(^{21}\). S100A8 is highly expressed in colorectal cancer and promotes epithelial–mesenchymal transition and metastasis via the transforming growth factor-\(\beta\)/USF2 axis\(^{22}\). In addition, another
study reported that by binding to NPTNβ, S100A8/9 activates NFIA and NFIB to induce the invasion and metastasis of lung cancer cells\cite{23}.

In this study, CIBERSORT analysis revealed that, compared with the mice in the normal group, the mice in the 4NQO group had a high abundance of M0 and M1 macrophages in their esophagus. Chronic inflammation plays an important role in the initiation and development of various cancers, particularly in digestive organs\cite{24}. Macrophages, as crucial components of innate immunity, are closely related to inflammatory and immune responses. During tumor initiation, macrophages form an inflammatory microenvironment that promotes mutation and growth. As tumors progress to malignancy, macrophages stimulate angiogenesis, enhance tumor cell migration and invasion, and suppress antitumor immunity\cite{25}. Numerous clinical and experimental studies have shown that macrophages promote cancer initiation and malignant progression. For example, in one meta-analysis, more than 80\% of studies reported a correlation between macrophage abundance and poor patient prognosis of patients\cite{26,27,28}. In ESCA, another study showed that after a surgical model of gastroesophageal reflux disease was established, numerous CD68+ macrophages surrounding squamous proliferative hyperplasia, Barrett’s metaplasia (BM), adenocarcinoma (AD), and SCC were identified, while moderate infiltration of CD163+ macrophages was observed in BM, AD, and SCC\cite{29}. These findings demonstrate the accretive effect of macrophages on esophageal carcinogenesis. Thus, macrophage infiltration induced by chronic inflammation may be involved in esophageal carcinogenesis.

Additionally, the study showed that inflammatory cytokines including IL-1β, IL-6, and IL-10 were dysregulated in peripheral blood. Previous studies have shown that transgenic mice that overexpress IL-1β in the esophagus can develop squamous epithelial atypia and subsequently develop SCC over time\cite{30}. Another study suggested that the overexpression of IL-1β was related to poor prognosis in patients with SCC and may be a promising molecular target for SCC therapeutic intervention\cite{31}. IL-6 has been well studied in SCC; for instance, positive IL-6 staining was linked to the development
of distant metastasis and decreased treatment response rates. In addition, the serum IL-6 concentration was significantly elevated in patients who developed disease failure. When IL-6 expression is inhibited, aggressive tumor behavior and radiation resistance can be overcome in vitro and in vivo\textsuperscript{32}. Another study revealed that polymorphism in IL-10 were more common found in patients with Barrett’s disease or AD and were previously reported to be associated with chronic inflammation\textsuperscript{33,34}. In summary, these results suggest that an imbalance in inflammation regulation could contribute to the development of ESCC.

The concept of multistage development of ESCC has been described previously. The strength of the present study is that we identified a group of genes that are possibly involved in esophageal tumorigenesis through transcriptome sequencing and bioinformatic analysis. Furthermore, the irregular infiltration of macrophages indicates chronic inflammation in the microenvironment of the esophagus, which may drive the process of carcinogenesis. Moreover, dysregulation of inflammation-related cytokines in the blood periphery implies a close correlation between inflammation and EC. Despite these strengths, our study inevitably has several limitations as well. First, because there is a lack of sequencing data focused on esophageal carcinogenesis, an ideal external dataset was not been found to validate the results of this study. Second, further functional experiments are required to investigate the biomarkers identified through transcriptome sequencing. In summary, subsequent validation in a clinical dataset and the use of additional experiments to validate the molecular function of the biomarkers will be the focus of our future research.

**CONCLUSION**

We induced esophageal carcinogenesis with 4NQO. HE staining confirmed the involvement of a multistage process in the development of ESCC. Then, we performed RNA-seq and utilized bioinformatics to identify key genes involved in the development of ESCC. The results revealed a group of key genes, and the functional enrichment analysis revealed that these genes were associated with keratinization, epidermal cell
differentiation, epidermal development, and epithelial cell proliferation. KEGG pathway analysis revealed inflammation-related and cancer-related signals, such as those related to IL-17. Cytokine-cytokine receptor interactions and transcriptional dysregulation may be involved in the development of EC. The CIBERSORT analysis revealed differences in the abundances of M0 and M1 macrophages between the 4NQO and normal groups. Furthermore, via IHC, we validated the results of transcriptome sequencing by measuring the expression of the two key genes (S100a8 and Krt6b) and macrophage markers.

**ARTICLE HIGHLIGHTS**

**Research background**

Precancerous lesions including low-grade intraepithelial neoplasia (LGIN) and high-grade IN (HGIN) are a mandatory stage in the development of esophageal squamous cell carcinoma (ESCC). The regulatory molecules driving esophageal carcinogenesis remain unclear, which largely hampers early esophageal cancer diagnosis and treatment.

**Research motivation**

We aimed to establish a mouse model mimicking the multi-stage development of ESCC and then to examine the key genes involved in the development of ESCC through transcriptome sequencing. Besides, we aimed to explore the immune cell infiltration in the microenvironment associated with the multi-stage progression of ESCC. Additionally, we hope to validate our results using human samples, which may provide a novel perspective for clinical application.

**Research objectives**

To examine the genes and immune cell infiltration in the microenvironment associated with the multi-stage progression of ESCC to facilitate diagnosis and early intervention.
Research methods
The mouse model of esophageal tumorigenesis was established by providing 4-nitroquinoline 1-oxide (4NQO) containing water to C57BL/6 mice. The transcriptome sequencing was performed for esophageal tissues with different pathological statuses. Differentially expressed gene analysis (DEGA) and Venn diagram were used to identify key genes. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were used to explore the biological function of genes. The CIBERSORT algorithm was used to detect the pattern of immune cell infiltration. The immunohistochemistry (IHC) was conducted to validate our results. Furthermore, the Luminex multiplex cytokines analysis was utilized to measure the serum level of cytokines in mice.

Research results
A total of 86 genes were identified as key genes in the development of ESCC. Enrichment analysis showed that these genes were mainly enriched in keratinization, epidermal cell differentiation, and interleukin (IL)-17 signaling pathways. CIBERSORT analysis showed that the esophageal carcinogenesis group had more infiltration of M0 and M1 macrophages which was then validated by IHC. Serum cytokines analysis showed that IL-1β and IL-6 were upregulated, while IL-10 was downregulated in LGIN, HGIN, and carcinoma groups. Moreover, the expression of the representative key genes, like S100a8 and Krt6b, was verified in external human samples, and the results of immunohistochemical staining were consistent with the findings in mice.

Research conclusions
We identified a set of key genes represented by S100a8 and Krt6b and investigated their potential biological functions. In addition, we found that macrophage infiltration and abnormal alterations of inflammation-associated cytokines, such as IL-1β, IL-6, and IL-10, in the peripheral blood may be closely associated with the development of ESCC.
Research perspectives

We have identified a set of key genes that are closely associated with esophageal carcinogenesis and may serve as novel biomarkers for early detection of ESCC. Additionally, the aberrant infiltration of inflammatory macrophages and dysregulation of cytokines may contribute to esophageal carcinogenesis.

Donghui Jin, Yousheng Mao. "Recent advances in the mechanisms of development and the early diagnosis and treatment of esophageal cancer", Holistic Integrative Oncology, 2023

Claire Wilson, Min Zi, Matthew Smith, Munir Hussain, Alicia D'Souza, Halina Dobrzynski, Mark R. Boyett. "Atrioventricular node dysfunction in pressure overload-induced heart failure—Involvement of the immune system and transcriptomic remodelling", Frontiers in Pharmacology, 2023