



# Weighted Correlation Network Analysis of Cancer Stem Cell-Related Prognostic Biomarkers in Esophageal Squamous Cell Carcinoma

Technology in Cancer Research & Treatment  
Volume 21: 1-9  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/15330338221117003  
journals.sagepub.com/home/tct  


Mengnan Zhao, PhD<sup>1,\*</sup>, Xing Jin, PhD<sup>1,\*</sup>, Zhencong Chen, PhD<sup>1,\*</sup>, Huan Zhang, PhD<sup>1</sup>, Cheng Zhan, PhD<sup>1</sup> , Hao Wang, PhD<sup>1</sup>, and Qun Wang, PhD<sup>1</sup>

## Abstract

**Background:** The role of cancer stem cells in esophageal squamous cell carcinoma (ESCC) remains unclear. **Methods:** The mRNA stemness index (mRNAsi) of 179 ESCC patients (GSE53625) was calculated using a machine learning algorithm based on their mRNA expression. Stemness-related genes were identified by weighted correlation network analysis (WGCNA) and LASSO regression, whose associations with mutation status, immune cell infiltrations, and potential compounds were also analyzed. The role of these genes in proliferation and their expressions was assessed in ESCC cell lines and 112 samples from our center. **Results:** The ESCC samples had significantly higher mRNAsi than the normal tissues. Patients with high mRNAsi exhibited higher worse OS. Seven stemness-related genes were identified by WGCNA and LASSO regression, based on which a risk-predicted score model was constructed. Among them, CST1, CILP, PITX2, F2RL2, and RIOX1 were favorable for OS, which were adverse for DPP4 and ZFH4 in the GSE53625 dataset. However, RIOX1 was unfavorable for OS in patients from our center. In vitro assays showed that CST1, CILP, PITX2, F2RL2, and RIOX1 were pro-proliferated, which were opposite for DPP4 and ZFH4. In addition, SMARCA4, NOTCH3, DNAH5, and KALRN were more mutated in the low-score group. The low-score group had significantly more memory B cells, monocytes, activated NK cells, and Tregs and less macrophages M2, resting mast cells, and resting dendritic cells. **Conclusions:** Seven stemness-related genes are significantly related to the prognosis, gene mutations, and immune cell infiltration of ESCC. Some potential anticancer compounds may be favorable for OS.

## Keywords

esophageal squamous cell carcinoma, WGCNA, GEO database, TCGA database, connectivity map database

## Abbreviations

ESCC, esophageal squamous cell carcinoma; CSC, cancer stem cell; mRNAsi, mRNA stemness index; WGCNA, weighted correlation network analysis; OS, overall survival; MoA, mechanisms of action.

Received: December 29, 2021; Revised: June 24, 2022; Accepted: July 12, 2022.

## Introduction

Esophageal squamous cell carcinoma (ESCC) occupied over 90% of all esophageal cancer cases, which exhibited a high incidence in East Asia and the Middle East. Due to exceptional tumor location, most ESCCs were found with dysphagia, which dramatically affects patient life quality. Unfortunately, the 5-year survival rate of esophageal cancer is <20% worldwide.<sup>1</sup> Thus, it is imperative to discover specific biomarkers to predict the prognosis and progress of ESCC.

Cancer stem cells (CSCs) also remain the capacity for self-renewal and differentiation, which are involved in tumor

<sup>1</sup> Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, Shanghai, China

\*These authors contributed equally to this work.

## Corresponding Authors:

Cheng Zhan, Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, 180 Fenglin Road, Xuhui District, Shanghai 200032, China.

Email: czhan10@fudan.edu.cn

Hao Wang, Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, 180 Fenglin Road, Xuhui District, Shanghai 200032, China.

Email: wang.hao@zs-hospital.sh.cn

Qun Wang, Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, 180 Fenglin Road, Xuhui District, Shanghai 200032, China.

Email: wang.qun@zs-hospital.sh.cn



progress, relapse, and treatment resistance.<sup>2</sup> First identified in breast cancer and leukemia, CSCs are now detected in various malignancies, including lung cancer, brain tumor, and intestinal tumor.<sup>2,3</sup>

The mRNA expression based-stemness index (mRNAsi) was first introduced by Malta et al<sup>4</sup> to assess the differentiation of cancer cells. Lian et al<sup>5</sup> reported that a higher mRNAsi was significantly unfavorable for OS in medulloblastoma patients. Weighted correlation network analysis (WGCNA) has been widely used to explore the highly correlated genes and their associations with specific parameters in high-throughput data with the advantages of improved integration and exactitude.<sup>6</sup>

In this research, we calculated the mRNAsi and identified the most relevant key genes with WGCNA and LASSO regression using profiles from the GEO database and TCGA database, based on which a risk-predicted score model was constructed and validated in ESCC patients from our center. Gene mutations and immune cell infiltrations in low- and high-score groups were also analyzed. Potential compounds targeting the score model were searched in the Connectivity Map database.

## Patients and Methods

### Ethics Approval

The Ethics Committee of Zhongshan Hospital, Fudan University has approved this research (Approval No. B2022-180R). Written informed consent was obtained from all patients.

### Data Collection and Calculation of mRNAsi

The mRNA-expression profiles and corresponding clinical information of 179 ESCC patients were obtained from the GSE53625 and 98 ESCC patients from the TCGA database.

The mRNAsi of each sample was calculated using the “TCGAbiolinks” R package.<sup>4,7,8</sup> The associations of mRNAsi with clinical information were assessed by the Wilcoxon rank-sum test owing to nonnormal distribution.

### Weighted Correlation Network Analysis and LASSO Regression

Weighted correlation network analysis was performed to screen coexpression modules and estimate their relevance to mRNAsi and other clinical parameters using the “WGCNA” R package. GO and KEGG pathway enrichment analysis was applied for genes within the most relevant modules to mRNAsi using the “clusterProfiler” R package.

The LASSO regression was utilized to screen ideal prognosis-related key genes within the most relevant to mRNAsi to mRNAsi by the “glmnet” R package.<sup>9</sup> The 1 – standard error (SE) criteria were applied to select the optimal model parameter  $\lambda$  as previously reported.<sup>10</sup>

### Construction and Validation of a Risk-Predicted Score Model

Univariate Cox regression was utilized to evaluate the prognostic value of key genes, based on which a risk-predicted score model was constructed. The risk score was the sum of the Cox regression coefficient ( $\beta$ ) multiplied by the corresponding gene expression. Patients were divided into low- and high-score groups according to the optimal cut-off score calculated using the “survminer” R package.<sup>11</sup>

The accuracy of the risk-score model was validated with 98 ESCC patients from the TCGA database. Immunofluorescence was used to estimate the expressions of the key genes in 112 ESCC samples from our center. Briefly, paraffin-embedded slides are deparaffinized and rehydrated. After antigen repair, the slides went through a block of endogenous peroxidase activity and nonspecific antigens and incubation with primary antibodies: anti-CST1 (Abcam, ab124281), anti-PITX2 (Affinity Biosciences LTD, DF13574), anti-RIOX1 (Abcam, ab194292), anti-CILP (Abcam, ab192881), anti-F2RL2 (Affinity Biosciences Ltd, DF15686), and anti-DPP4 (Affinity Biosciences Ltd, DF12387), anti-ZFH4 (Affinity Biosciences Ltd, DF10016), horseradish peroxidase-conjugated secondary antibody, and Opal tyramide signal amplification (TSA) Fluorochromes (Opal 2-Color Manual IHC Kit, G1226, Servicebio Co., Ltd) for 10 min at room temperature. After the second run, the slides were stained with DAPI. The ImageJ software was used to analyze the fluorescence intensity.<sup>12</sup> The effects of their expressions on prognosis were assessed by Cox regression.

### Cell Culture and siRNA Transfection

KYSE150 (ESCC cell lines) were cultured with RPMI-1640 medium (KeyGEN BioTECH), supplemented with 10% fetal bovine serum (Every Green) and 100 IU/mL penicillin/streptomycin (Beyotime) in a humidified 5% CO<sub>2</sub> atmosphere at 37 °C.

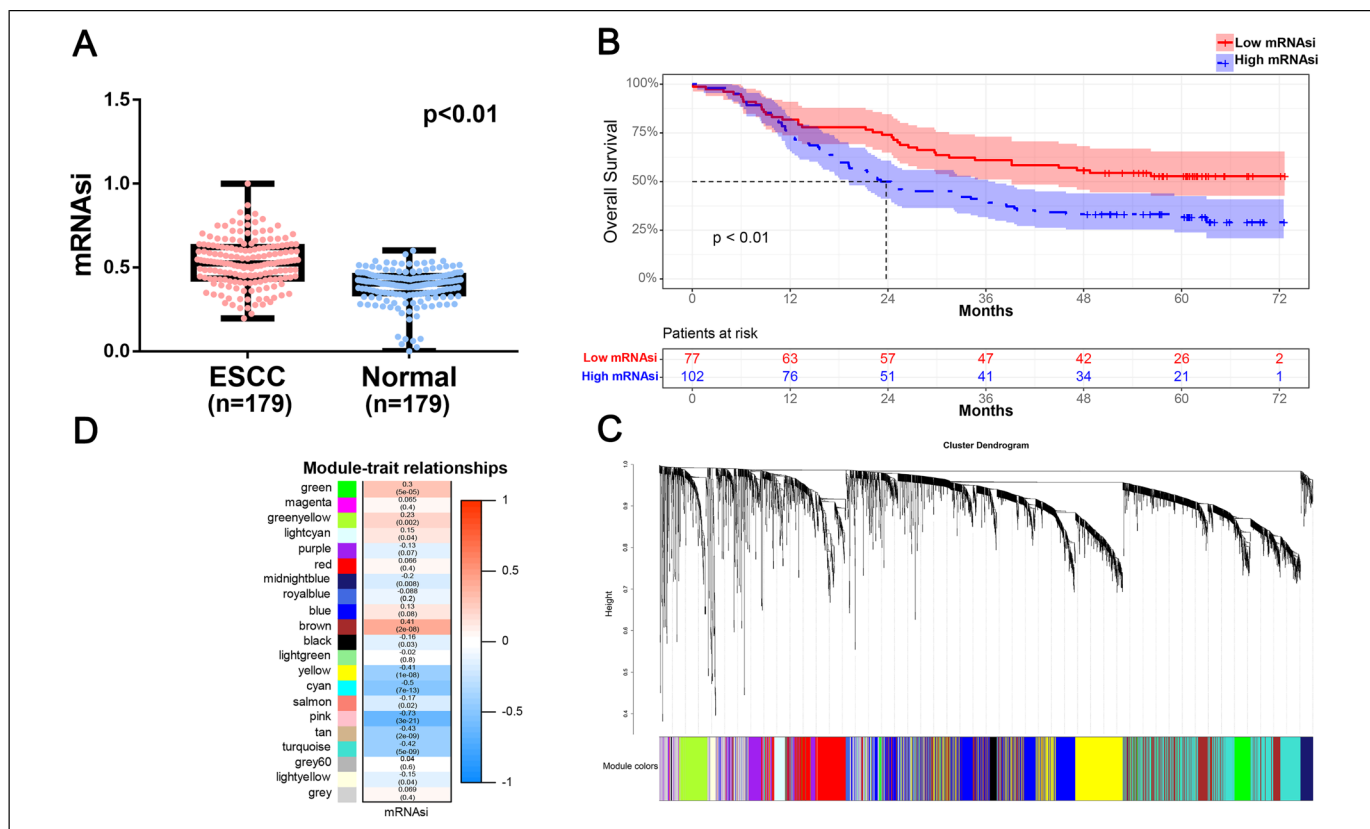
Two different siRNAs were transfected into ESCC cells to knock down gene expressions with Lipo8000 transfect reagent (Beyotime) and Opti-MEM (Thermo Fisher Scientific) according to the manufacturer’s protocol (Table S2).

### RNA Extraction and Quantitative Real-Time Polymerase Chain Reaction

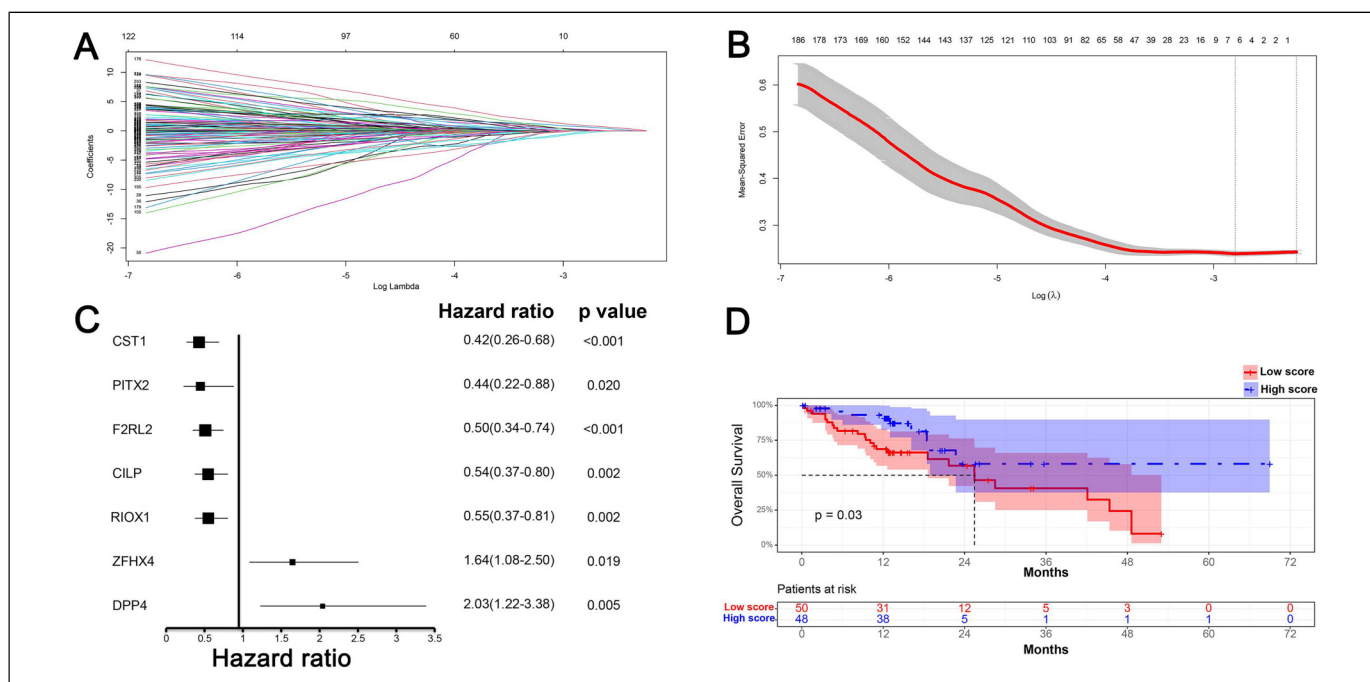
Total RNA was obtained with RNAiso Plus (Takara Bio, 9108) according to the manufacturer’s protocol. RT reagent Kit (Takara Bio, RR047A) was used to synthesize the cDNA template. A quantitative real-time-polymerase chain reaction was performed using TB Green (Takara Biomedical Technology, RR820A) in QuantStudio5 (Thermo Fisher Scientific) according to the manufacturer’s protocol. Primer sequences are listed in Table S2.

### Cell Proliferation Assays

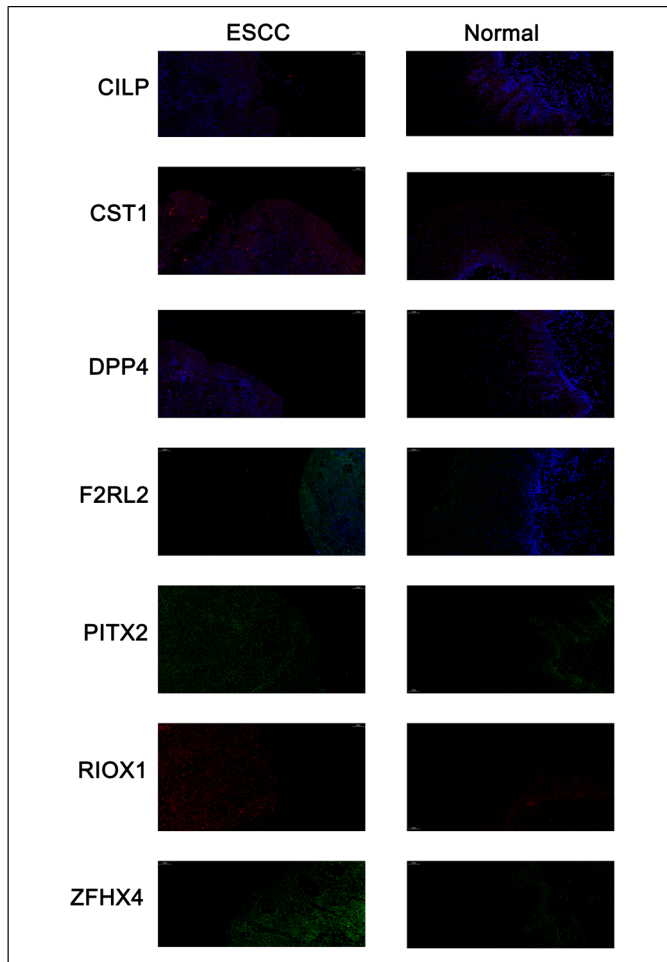
Briefly, 2000 cells/well were inoculated in blank 96-well plates. After incubation for 24, 48, 72, 96, and 120 h, cell proliferation was measured by CCK-8 (Beyotime).



**Figure 1.** (A) mRNAi of ESCC and normal esophagus, (B) overall survival between low and high mRNAi group, (C) 21 modules generated by the WGCNA algorithm, and (D) correlations between 21 modules and mRNAi in GES53625 patients. Abbreviations: mRNAi, mRNA stemness index; ESCC, esophageal squamous cell carcinoma; WGCNA, weighted correlation network analysis.



**Figure 2.** (A) Fitting curve of coefficients and lambda, (B) relationships between partial likelihood deviation and log lambda in LASSO regression, (C) Forest map of mRNA expressions of 7 prognosis-related key genes in GES53625 dataset, (D) Overall survival between low and high score group in TCGA esophageal squamous cell carcinoma (ESCC) patients.



**Figure 3.** Representative immunofluorescence of CILP (red), CST1 (red), DPP4 (red), F2RL2 (green), PITX2 (green), RIOX1 (red), ZFHX4 (green) in esophageal squamous cell carcinoma (ESCC) samples and normal tissues of 112 ESCC patients from Zhongshan Hospital, Fudan University.

### Characteristics of Mutations and Immune Cell Infiltration

Various mutations and correlations between different gene mutations were analyzed with the “maftools” R package.<sup>13,14</sup> The relative abundance of the 22 immune cell subpopulations was assessed by the CIBERSORT method.<sup>15</sup>

### Potential Anticancer Compounds

The potential activated or inhibited compounds were screened in the Connectivity Map (CMap) database.<sup>16</sup>

### Statistical Analysis

All statistical analyses were performed with SPSS (version 24; IBM). All tests were two-sided and  $P < .05$  was considered significant.

## Results

### Calculation of mRNAsi and Weighted Correlation Network Analysis of Most Relevant Module

Based on the mRNA expression, we calculated the mRNAsi of 179 pairs of samples from ESCC patients in the GSE53625 dataset. Obviously, the ESCC samples had significantly higher mRNAsi than the normal tissues ( $P < .01$ ) (Figure 1A). No significant association of mRNAsi with age, sex, and stage was observed. However, poorly differentiated cases and smokers had a higher mRNAsi (Figure S1A). The OS was also significantly worse in the patients with high mRNAsi ( $P < .01$ ) (Figure 1B).

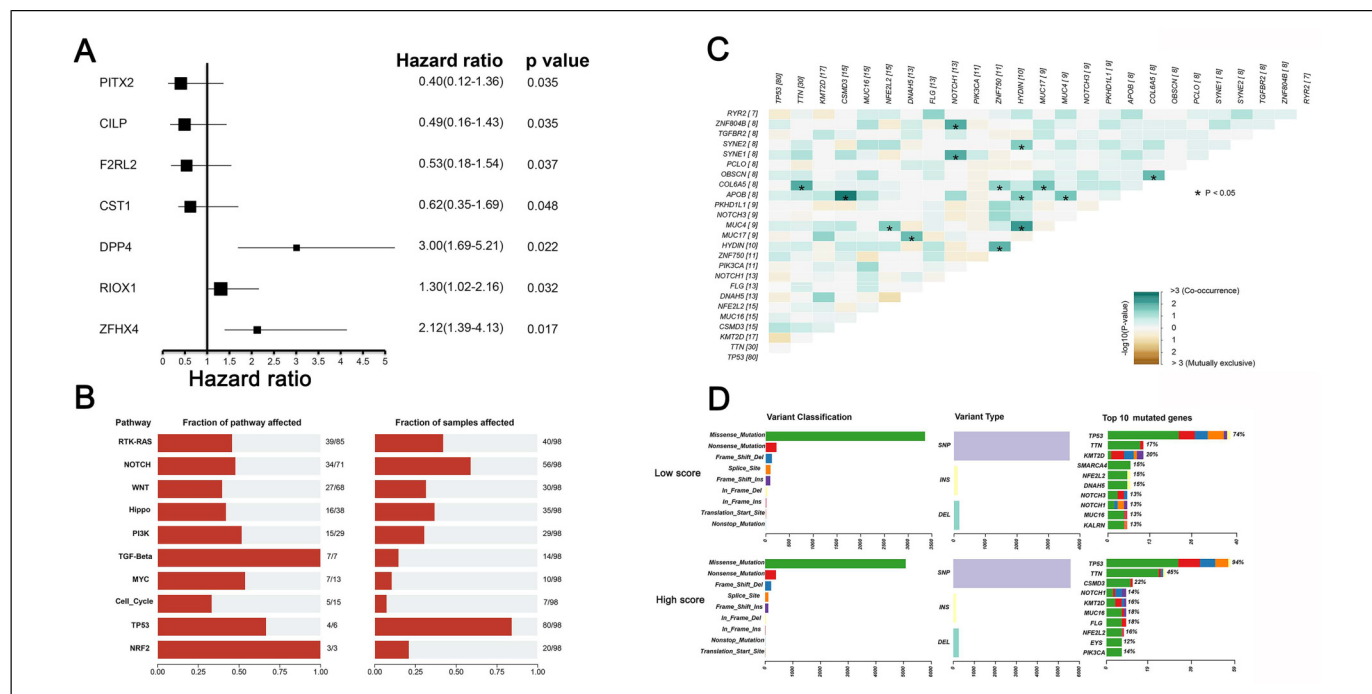
Of the 21 gene modules generated by WGCNA, the pink module was the most relevant to the mRNAsi with a correlation coefficient of  $-0.73$  ( $P < .01$ ) (Figure 1C). Figure S2A also revealed that genes in the pink module were apparently related to the mRNAsi. GO and KEGG pathway enrichment analysis showed that extracellular matrix constituent and structure organization, PI3K-Akt signaling pathway, Rap1 signaling pathway, Rap1 signaling pathway, and calcium signaling pathway were enriched in the pink module (Figure 1D).

### Construction and Validation of a Risk-Predicted Score Model

With the LASSO regression, 7 prognosis-related key genes among genes in the pink module were identified, based on which a risk-predicted score model was constructed with their Cox regression coefficient (Figure 2A-C). Risk score =  $(0.42 \times \text{expression level of CST1}) + (0.44 \times \text{expression level of PITX2}) + (0.50 \times \text{expression level of F2RL2}) + (0.54 \times \text{expression level of CILP}) + (0.55 \times \text{expression level of RIOX1}) + (1.64 \times \text{expression level of ZFHX4}) + (2.03 \times \text{expression level of DPP4})$ . No significant association of risk scores with age, smoke, and stage was observed. However, poorly differentiated ( $P < .01$ ) and male cases ( $P = .06$ ) showed significantly lower risk scores (Figure S1B). Subsequently, 179 patients were divided into high- and low-score groups according to the optimal cut-off score calculated by the “survminer” R package. The low-score group exhibited a significantly worse OS ( $P = .04$ ), which was also observed in the 98 ESCC patients from the TCGA database ( $P = .03$ ) (Figure 2D).

Besides, the expressions of the 7 key genes were also measured by immunofluorescence in 112 ESCC patients from our center (Table S1). CST1, PITX2, RIOX1, F2RL2, ZFHX4, and DPP4 were up-regulated in the ESCC samples. Conversely, CILP was down-regulated in the ESCC samples (Figure 3). In addition, high expressions of CST1 ( $P = .048$ ), PITX2 ( $P = .035$ ), CILP ( $P = .035$ ), and F2RL2 ( $P = .037$ ) were favorable for survival, which were adverse for DPP4 ( $P = .022$ ), RIOX1 ( $P = .032$ ), and ZFHX4 ( $P = .017$ ) (Figure 4A).

Through down-regulating the CST1, CILP, PITX2, F2RL2, and RIOX1 expressions with siRNAs, we observed that the proliferation of KYSE150 cells was inhibited, which was the opposite for DPP4 and ZFHX4 (Figure S2B).



**Figure 4.** (A) Forest map of fluorescence intensity of 7 prognosis-related key genes in 112 esophageal squamous cell carcinoma (ESCC) patients from Zhongshan Hospital, Fudan University, (B) fraction of affected pathway and samples in TCGA ESCC patients, (C) correlation map of top 25 mutated genes in ESCC from the TCGA database, and (D) gene mutation characteristics between low- and high-score group in TCGA ESCC patients.

### Genetic Variations and Immune Cell Infiltration

The P53, NOTCH, and RKT-RAS pathways were the most common mutated pathways in ESCC (Figure 4B). Interestingly, COL6A5 was likely to comutate with TTN, ZNF750, and MUC17 ( $P < .05$ , respectively). Cooccurrence of APOB and CSMD3, HYDIN, and MUC4 was also observed ( $P < .05$ , respectively) (Figure 4C).

Missense mutation and SNP were the most frequent in both low- and high-score groups. More specifically, TP53, TTN, CSMD3, FLG, EYS, and PIK3CA were less mutated in the low-score group. However, SMARCA4, NOTCH3, DNAH5, and KALRN were more mutated in the low-score group (Figure 4D).

The CIBERSORT algorithm showed that the low-score group had significantly more memory B cells, monocytes, activated NK cells, and Tregs and less macrophages M2, resting mast cells, and resting dendritic cells ( $P < .05$ , respectively) (Figure 5A).

### Screening of Potential Targeted Compounds

Three hundred top differentially expressed genes between low- and high-score groups were queried in the CMap database. Finally, 44 compounds were identified, which were related to 36 mechanisms of action (MoA). Among them, PDGFR receptor inhibitor, VEGFR inhibitor, KIT inhibitor, potassium channel activator, and FLT3 inhibitor were shared frequently (Figure 5B).

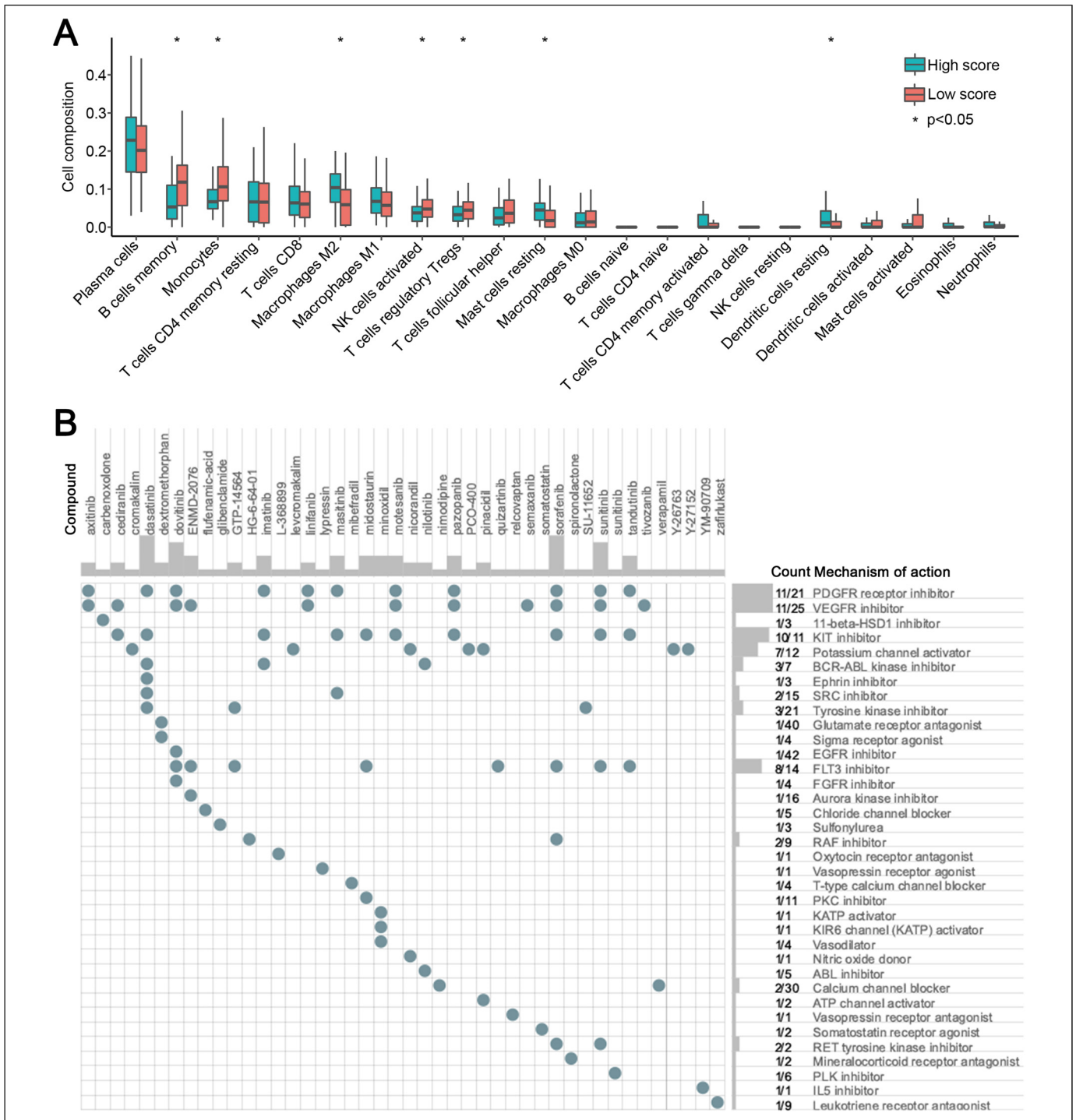
### Discussion

Despite great progress in examination techniques and therapy, most ESCC patients reach an advanced stage, and lose the chance for radical surgery. Moreover, quite a few patients have recurrent or metastatic ESCC. A comprehensive understanding of the molecular heterogeneity of ESCC could improve personalized therapies.

It has been reported that basal cells are the progenitor cells of the esophageal squamous epithelium.<sup>17,18</sup> Stimulated by interleukin-1 $\beta$  (IL-1 $\beta$ )/interleukin-6 (IL-6), the basal cells can be transformed, accounting for ESCC.<sup>19</sup> Many research studies have proven the existence of CSCs in ESCC.<sup>20–23</sup> Besides, many cell surface markers of CSCs in ESCC have been discovered including CD271, CD44, CD90, CD133, and CXCR4, which were unfavorable for prognosis and able to predict and evaluate therapy response.<sup>24–32</sup> We also observed that low differentiated grades meant poor OS for ESCC patients.

Secretory cystatin SN encoding by CST1 belongs to the type 2 cystatin superfamily, which specifically inhibits the proteolytic activity of cysteine proteases.<sup>33</sup> Choi et al<sup>34</sup> reported that CST1 was up-regulated in gastric cancer and contributed to cancer cell proliferation, which was also observed in colorectal cancer.<sup>35</sup> Similar to their findings, ESCC exhibited upregulation of CST1 and facilitation of cell proliferation in our study. PITX2 belongs to the bicoid/paired-like homeobox gene family with vital roles in embryonic development.<sup>36</sup> It has been reported that the expression of PITX2 was increased in ovarian cancer and colorectal cancer.<sup>37,38</sup> It promoted the invasion of ovarian





**Figure 5.** (A) The abundance of 22 immune cell subpopulations in the low and high score group based on the CIBERSORT algorithm and (B) potential inhibitors targeting the risk model based on the CMap database.

cancer. However, Hirose et al<sup>38</sup> revealed its inhibition of cell growth and invasion in colorectal cancer. We exhibited its over-expression and facilitation of cell proliferation in ESCC, which need further research to solve the contradiction. F2RL2, also known as PAR3, is a polarity protein, which regulates apical/basal polarity and spindle orientation and is essential for stem cell maintenance.<sup>39</sup> Zhou et al<sup>40</sup> reported that loss of Par3

promoted prostatic tumorigenesis. Nevertheless, Dadras et al<sup>41</sup> showed that Par3 participated in homeostatic redox control and limited invasiveness of glioblastoma.

CILP is also a secreted protein, which is mainly from articular cartilage chondrocytes.<sup>42</sup> Its role in malignancy is rarely reported. We found that it was down-regulated in ESCC. Interestingly, its high expression in ESCC was related to

worse OS and promotion of ESCC cells, which need larger sample size research studies to claim its effect. RIOX1 is one of the JMJD histone demethylases and regulates target genes in embryonic stem cells.<sup>43,44</sup> Sinha et al<sup>45</sup> demonstrated that RIOX1 was highly expressed in prostate cancer and promoted cancer cell survival. It has been reported that ZFH4 was over-expressed in ESCC and facilitated cancer migration and invasion,<sup>46</sup> which was similar to our results. DPP4 is a type II transmembrane protein, whose soluble form can be easily detected in serum or plasma. There are research studies that it was down-regulated in melanoma, nonsmall cell lung cancer, and renal cell carcinoma,<sup>47–49</sup> up-regulated in thyroid cancer, ovarian cancer, and ESCC.<sup>50–52</sup> Goscinski et al<sup>53</sup> revealed that high expression of DPP4 correlated with better survival in ESCC patients, which was also observed in our study.

Many research studies have reported that NOTCH3 was differentially expressed between malignant and corresponding normal tissues, which was an important prognostic factor.<sup>54</sup> Liu et al<sup>55</sup> found that NOTCH3 was up-regulated in ovarian epithelial cancer than in normal tissues and in an ovarian benign tumor, which meant a shorter OS. However, Zhou et al<sup>56</sup> reported that NOTCH3 was down-regulated in small cell lung cancer. In addition, NOTCH3 mutations were closely associated with Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL) and Pulmonary Hypertension.<sup>57</sup> Sawada et al<sup>58</sup> demonstrated that NOTCH3 was frequently mutated (7.6%) in ESCC. In our study, NOTCH3 mutation was related to the poor OS of ESCC patients. Mangalparthi et al<sup>59</sup> reported that DNAH5 was frequently mutated (21%) in ESCC, which meant poor survival and was consistent with our study. KARLN was mutated in a variety of malignancies, including melanoma, glioblastoma multiforme, lung cancer, and colorectal cancer.<sup>60</sup> Grønhøj et al<sup>61</sup> reported that KALRN was frequently mutated (28%) in human papillomavirus positive (HPV+) oropharyngeal squamous cell carcinomas.

Dasatinib is an oral tyrosine kinase inhibitor and has been widely used in CML and Philadelphia chromosome-positive acute lymphoblastic leukemia.<sup>62</sup> Additionally, several clinical trials of dasatinib as monotherapy or combined therapy with other drugs have exhibited promising antitumor effects in solid tumors, including breast cancer, nonsmall cell lung cancer, melanoma, and colorectal cancer.<sup>63</sup> Sorafenib is an oral multikinase inhibitor and was first approved for the treatment of renal cell cancer. Xiang et al<sup>64</sup> demonstrated that stemness contributed to sorafenib resistance. Several clinical trials with dasatinib in nonsmall cell lung cancer and prostate cancer also showed an anticancer effect.<sup>65</sup> The association of these compounds with ESCC stemness and prognosis requires further study.

Inevitably, there are some limitations to this research. First, this study is retrospective. Second, because of the limited sample size, the power calculation was not done for its estimation. Third, the effect of key genes on ESCC was only explored with one cell line. More experiments need to be done to validate it, including proliferation, colony formation, migration, and animal experiments.

## Conclusion

In conclusion, high mRNAsi signified a poor OS in ESCC patients. Seven stemness-related genes were identified to be significantly associated with OS by WGCNA and LASSO regression, which were validated in 112 ESCC patients from our center and in vitro experiments. Two groups with different expression patterns exhibited distinct mutation characteristics and immune cell infiltration. The potential mechanisms between stemness and tumor immunity and drug sensitivity still need further studies.

## Acknowledgments

We would like to thank International Science Editing Co. for the language editing service.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Shanghai Pujiang Program (grant number 2020PJD009).

## Ethics

This study was approved by the Ethics Committee of Zhongshan Hospital, Fudan University, Shanghai, China (Approval No. B2017-153). Written informed consent was obtained from all patients.

## ORCID iD

Cheng Zhan  <https://orcid.org/0000-0001-8745-9276>

## Supplemental Material

Supplementary material for this article is available online.

## References

1. Smyth EC, Lagergren J, Fitzgerald RC, et al. Oesophageal cancer. *Nat Rev Dis Primers*. 2017;3(1):17048.
2. Singh S, Chellappan S. Lung cancer stem cells: molecular features and therapeutic targets. *Mol Aspects Med*. 2014;39:50-60.
3. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*. 1997;3(7):730-737.
4. Malta TM, Sokolov A, Gentles AJ, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell*. 2018;173(2):338-354.e15.
5. Lian H, Han YP, Zhang YC, et al. Integrative analysis of gene expression and DNA methylation through one-class logistic regression machine learning identifies stemness features in medulloblastoma. *Mol Oncol*. 2019;13(10):2227-2245.
6. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf*. 2008;9(1):559.

7. Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016;44(8):e71.
8. Mounir M, Lucchetta M, Silva TC, et al. New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput Biol.* 2019;15(3):e1006701.
9. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* 1997;16(4):385-395.
10. Jiang Y, Zhang Q, Hu Y, et al. Immunoscore signature: a prognostic and predictive tool in gastric cancer. *Ann Surg.* 2018;267(3):504-513.
11. Liang JY, Wang DS, Lin HC, et al. A novel ferroptosis-related gene signature for overall survival prediction in patients with hepatocellular carcinoma. *Int J Biol Sci.* 2020;16(13):2430-2441.
12. Jensen EC. Quantitative analysis of histological staining and fluorescence using ImageJ. *Anat Rec.* 2013;296(3):378-381.
13. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 2018;28(11):1747-1756.
14. Sha D, Jin Z, Budczies J, Kluck K, Stenzinger A, Sinicrope FA. Tumor mutational burden as a predictive biomarker in solid tumors. *Cancer Discov.* 2020;10(12):1808-1825.
15. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12(5):453-457.
16. Subramanian A, Narayan R, Corsello SM, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell.* 2017;171(6):1437-1452.e17.
17. Zhang JX, Chen ZH, Xu Y, et al. Downregulation of MicroRNA-644a promotes esophageal squamous cell carcinoma aggressiveness and stem cell-like phenotype via dysregulation of PITX2. *Clin Cancer Res.* 2017;23(1):298-310.
18. Wang X, Ouyang H, Yamamoto Y, et al. Residual embryonic cells as precursors of a Barrett's-like metaplasia. *Cell.* 2011;145(7):1023-1035.
19. Liu K, Zhao T, Wang J, et al. Etiology, cancer stem cells and potential diagnostic biomarkers for esophageal cancer. *Cancer Lett.* 2019;458:21-28.
20. Shiozaki A, Kudou M, Ichikawa D, et al. Esophageal cancer stem cells are suppressed by tranilast, a TRPV2 channel inhibitor. *J Gastroenterol.* 2018;53(2):197-207.
21. Liu Q, Cui X, Yu X, et al. Cripto-1 acts as a functional marker of cancer stem-like cells and predicts prognosis of the patients in esophageal squamous cell carcinoma. *Mol Cancer.* 2017;16(1):81.
22. Yuan X, Kong J, Ma Z, et al. KDM4C, a H3K9me3 histone demethylase, is involved in the maintenance of human ESCC-initiating cells by epigenetically enhancing SOX2 expression. *Neoplasia.* 2016;18(10):594-609.
23. Islam F, Gopalan V, Wahab R, Smith RA, Lam AK. Cancer stem cells in oesophageal squamous cell carcinoma: identification, prognostic and treatment perspectives. *Crit Rev Oncol Hematol.* 2015;96(1):9-19.
24. Huang SD, Yuan Y, Liu XH, et al. Self-renewal and chemotherapy resistance of p75NTR positive cells in esophageal squamous cell carcinomas. *BMC Cancer.* 2009;9(1):9.
25. Yamaguchi T, Okumura T, Hirano K, et al. P75 neurotrophin receptor expression is a characteristic of the mitotically quiescent cancer stem cell population present in esophageal squamous cell carcinoma. *Int J Oncol.* 2016;48(5):1943-1954.
26. Okumura T, Tsunoda S, Mori Y, et al. The biological role of the low-affinity p75 neurotrophin receptor in esophageal squamous cell carcinoma. *Clin Cancer Res.* 2006;12(17):5096-5103.
27. Zhao JS, Li WJ, Ge D, et al. Tumor initiating cells in esophageal squamous cell carcinomas express high levels of CD44. *PLoS One.* 2011;6(6):e21419.
28. Tang KH, Dai YD, Tong M, et al. A CD90(+) tumor-initiating cell population with an aggressive signature and metastatic capacity in esophageal cancer. *Cancer Res.* 2013;73(7):2322-2332.
29. Lu C, Xu F, Gu J, et al. Clinical and biological significance of stem-like CD133(+)CXCR4(+) cells in esophageal squamous cell carcinoma. *J Thorac Cardiovasc Surg.* 2015;150(2):386-395.
30. Wang Y, Zhe H, Gao P, Zhang N, Li G, Qin J. Cancer stem cell marker ALDH1 expression is associated with lymph node metastasis and poor survival in esophageal squamous cell carcinoma: a study from high incidence area of northern China. *Dis Esophagus.* 2012;25(6):560-565.
31. Hang D, Dong HC, Ning T, Dong B, Hou DL, Xu WG Prognostic value of the stem cell markers CD133 and ABCG2 expression in esophageal squamous cell carcinoma. *Dis Esophagus.* 2012;25(7):638-644.
32. Smit JK, Faber H, Niemantsverdriet M, et al. Prediction of response to radiotherapy in the treatment of esophageal cancer using stem cell markers. *Radiother Oncol.* 2013;107(3):434-441.
33. Lah TT, Babnik J, Schiffmann E, Turk V, Skaleric U. Cysteine proteinases and inhibitors in inflammation: their role in periodontal disease. *J Periodontol.* 1993;64(Suppl 5):485-491.
34. Choi EH, Kim JT, Kim JH, et al. Upregulation of the cysteine protease inhibitor, cystatin SN, contributes to cell proliferation and cathepsin inhibition in gastric cancer. *Clin Chim Acta.* 2009;406(1-2):45-51.
35. Kim JT, Lee SJ, Kang MA, et al. Cystatin SN neutralizes the inhibitory effect of cystatin C on cathepsin B activity. *Cell Death Dis.* 2013;4(12):e974.
36. Lee WK, Thévenod F. Oncogenic PITX2 facilitates tumor cell drug resistance by inverse regulation of hOCT3/SLC22A3 and ABC drug transporters in colon and kidney cancers. *Cancer Lett.* 2019;449:237-251.
37. Basu M, Bhattacharya R, Ray U, Mukhopadhyay S, Chatterjee U, Roy SS. Invasion of ovarian cancer cells is induced by PITX2-mediated activation of TGF- $\beta$  and Activin-A. *Mol Cancer.* 2015;14(1):162.
38. Hirose H, Ishii H, Mimori K, et al. The significance of PITX2 overexpression in human colorectal cancer. *Ann Surg Oncol.* 2011;18(10):3005-3012.
39. Huo Y, Macara IG. The Par3-like polarity protein Par3L is essential for mammary stem cell maintenance. *Nat Cell Biol.* 2014;16(6):529-537.
40. Zhou PJ, Wang X, An N, et al. Loss of Par3 promotes prostatic tumorigenesis by enhancing cell growth and changing cell division modes. *Oncogene.* 2019;38(12):2192-2205.
41. Dadras MS, Caja L, Mezheyski A, et al. The polarity protein Par3 coordinates positively self-renewal and



- negatively invasiveness in glioblastoma. *Cell Death Dis.* 2021;12(10):932.
42. Lorenzo P, Bayliss MT, Heinegård D. A novel cartilage protein (CILP) present in the mid-zone of human articular cartilage increases with age. *J Biol Chem.* 1998;273(36):23463-8.
  43. Højfeldt JW, Agger K, Helin K. Histone lysine demethylases as targets for anticancer therapy. *Nat Rev Drug Discov.* 2013;12(12):917-930.
  44. Nishizawa Y, Nishida N, Konno M, et al. Clinical significance of histone demethylase NO66 in invasive colorectal cancer. *Ann Surg Oncol.* 2017;24(3):841-849.
  45. Sinha KM, Bagheri-Yarmand R, Lahiri S, et al. Oncogenic and osteolytic functions of histone demethylase NO66 in castration-resistant prostate cancer. *Oncogene.* 2019;38(25):5038-5049.
  46. Qing T, Zhu S, Suo C, Zhang L, Zheng Y, Shi L. Somatic mutations in ZFH4 gene are associated with poor overall survival of Chinese esophageal squamous cell carcinoma patients. *Sci Rep.* 2017;7(1):4951.
  47. Enz N, Vliegen G, De Meester I, Jungraithmayr W. CD26/DPP4 – a potential biomarker and target for cancer therapy. *Pharmacol Ther.* 2019;198:135-159.
  48. Wesley UV, Albino AP, Tiwari S, Houghton AN. A role for dipeptidyl peptidase IV in suppressing the malignant phenotype of melanocytic cells. *J Exp Med.* 1999;190(3):311-322.
  49. Wesley UV, Tiwari S, Houghton AN. Role for dipeptidyl peptidase IV in tumor suppression of human non small cell lung carcinoma cells. *Int J Cancer.* 2004;109(6):855-866.
  50. Kotani T, Aratake Y, Ogata Y, et al. Expression of dipeptidyl aminopeptidase IV activity in thyroid carcinoma. *Cancer Lett.* 1991;57(3):203-208.
  51. Zhang M, Xu L, Wang X, Sun B, Ding J. Expression levels of seprase/FAP $\alpha$  and DPPIV/CD26 in epithelial ovarian carcinoma. *Oncol Lett.* 2015;10(1):34-42.
  52. Augoff K, Hryniewicz-Jankowska A, Tabola R, et al. Upregulated expression and activation of membrane-associated proteases in esophageal squamous cell carcinoma. *Oncol Rep.* 2014;31(6):2820-2826.
  53. Goscinski MA, Suo ZH, Nesland JM, et al. Seprase, dipeptidyl peptidase IV and urokinase-type plasminogen activator expression in dysplasia and invasive squamous cell carcinoma of the esophagus. A study of 229 cases from Anyang Tumor Hospital, Henan Province, China. *Oncology.* 2008;75(1-2):49-59.
  54. Aburjania Z, Jang S, Whitt J, Jaskula-Stzul R, Chen H, Rose JB. The role of Notch3 in cancer. *Oncologist.* 2018;23(8):900-911.
  55. Liu Z, Yun R, Yu X, et al. Overexpression of Notch3 and pS6 is associated with poor prognosis in human ovarian epithelial cancer. *Mediators Inflamm.* 2016;2016:5953498.
  56. Zhou M, Jin WY, Fan ZW, Han RC. Analysis of the expression of the Notch3 receptor protein in adult lung cancer. *Oncol Lett.* 2013;5(2):499-504.
  57. Hosseini-Alghaderi S, Baron M. Notch3 in development, health and disease. *Biomolecules.* 2020;10(3):485.
  58. Sawada G, Niida A, Uchi R, et al. Genomic landscape of esophageal squamous cell carcinoma in a Japanese population. *Gastroenterology.* 2016;150(5):1171-1182.
  59. Mangalparthi KK, Patel K, Khan AA, et al. Mutational landscape of esophageal squamous cell carcinoma in an Indian cohort. *Front Oncol.* 2020;10:1457.
  60. Tate JG, Bamford S, Jubb HC, et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* 2019;47(D1):D941-D947.
  61. Grønhoj C, Jensen DH, Agander T, et al. Deep sequencing of human papillomavirus positive loco-regionally advanced oropharyngeal squamous cell carcinomas reveals novel mutational signature. *BMC Cancer.* 2018;18(1):640.
  62. Brave M, Goodman V, Kaminskas E, et al. Sprycel for chronic myeloid leukemia and Philadelphia chromosome-positive acute lymphoblastic leukemia resistant to or intolerant of imatinib mesylate. *Clin Cancer Res.* 2008;14(2):352-359.
  63. Araujo J, Logothetis C. Dasatinib: a potent SRC inhibitor in clinical development for the treatment of solid tumors. *Cancer Treat Rev.* 2010;36(6):492-500.
  64. Xiang DM, Sun W, Zhou T, et al. Oncofetal HLF transactivates c-Jun to promote hepatocellular carcinoma development and sorafenib resistance. *Gut.* 2019;68(10):1858-1871.
  65. Hahn O, Stadler W. Sorafenib. *Curr Opin Oncol.* 2006;18(6):615-621.