Am J Cancer Res 2024;14(5):2343-2370
www.ajcr.us /ISSN:2156-6976/ajcr0155040

Original Article

ANOS1 accelerates the progression of esophageal cancer identified by multi-omic approaches

Zuoquan Zhu[1], Shikun Dong[2], Shaolei Qin[3], Ke Gu[1], Yanjun Zhou[1]

[1]Department of Radiotherapy and Oncology, The Affiliated Hospital of Jiangnan University, Wuxi 214000, Jiangsu, China; [2]Department of Otorhinolaryngology, Zhongda Hospital, Southeast University, Nanjing 210009, Jiangsu, China; [3]Jiangnan University, Wuxi 214000, Jiangsu, China

Received December 19, 2023; Accepted April 27, 2024; Epub May 15, 2024; Published May 30, 2024

Abstract: To assess the role of ANOS1 in esophageal cancer (ESCA) progression, multi-omic analysis and experimental validation were employed. It was revealed that ANOS1 expression is significantly enhanced in ESCA patients and cell lines. The expression level of ANOS1 in ESCA patients can distinguish the malignancy from normal tissue with an area under curve (AUC) >0.75. Moreover, increased expression of ANOS1 is associated with advanced T stage and worse disease-free survival of ESCA patients. Therefore, a clinically applicable nomogram with ANOS1 was established with strong predictive power. Furthermore, high expression of ANOS1 in ESCA is correlated with (i) the enrichment of epithelial-mesenchymal transition by gene set enrichment analysis, (ii) the involvement in hypoxia, angiogenesis, WNT signaling pathway, and TGFβ signaling pathway by gene set variation analysis, (iii) the presence of the small insertion and deletion mutational signature ID9, associated with chromothripsis, in the single-nucleotide polymorphism analysis, (iv) the amplification of 11q13.3 in the copy number variants analysis, (v) the enrichment of cancer-associated fibroblasts and mesenchymal stromal cells in the tumor microenvironment. All the results from multi-omic analysis indicate that ANOS1 plays a pivotal role in accelerating the progression of ESCA. Results from *in vivo* and *in vitro* experiments show that the knockdown of ANOS1 hampers the proliferation of ESCA cells, further validating the oncogenic role of ANOS1 in ESCA. Additionally, potential chemotherapeutics with sensitivity were identified in the high-ANOS1 group. In conclusion, ANOS1 accelerates the progression of ESCA.

Keywords: ANOS1, esophageal cancer, proliferation

Introduction

Esophageal cancer (ESCA) is one of the invasive tumors, ranking seventh and sixth in global morbidity and mortality, respectively. Over 70% of esophageal cancer patients have esophageal squamous cell carcinoma (ESCC), and the incidence of ESCC is particularly high in some parts of East Asia and Africa [1]. Due to the difficulty of early diagnosis, most ESCA patients are diagnosed at a locally advanced stage. The majority of patients experience recurrence and metastasis, resulting in a 5-year survival rate between 10% and 33% [2]. Therefore, early diagnosis and treatment of esophageal cancer are crucial for improving patients' quality of life and survival rates.

The ANOS1 gene encodes anosmin-1, an extracellular matrix-associated glycosylated protein. Anosmin-1 consists of an N-terminal cysteine-rich domain, a whey acidic protein (WAP) domain, evolutionarily conserved fibronectin type III domains, and a histidine-rich C-terminal region [3]. ANOS1 promotes neurite outgrowth and cell migration through the fibroblast growth factor receptor 1 signaling pathway [4]. ANOS1 enhances the invasion of brain malignancy by modulating cell adhesion, activating extracellular proteases *in vitro*, and promoting the proliferation of cancer cells [5]. Anosmin-1 facilitates colon cancer cell migration and exhibits an anti-apoptotic capacity [6], which demonstrates the multifunction of ANOS1 in neurodevelopment and cancer progression. However, it is worth noting that ANOS1 has also been reported as a suppressor for hepatocellular cancer [7].

Although numerous studies have focused on the role of anosmin-1, its function in ESCA

https://doi.org/10.62347/SPCP3536

remains unclear. In this study, we investigated the potential role of ANOS1 in ESCA progression for the first time. Our findings demonstrate that ANOS1 is commonly overexpressed in ESCA patients and cell lines, and its expression level can effectively distinguish malignant from normal tissue. Moreover, high ANOS1 expression is associated with advanced T stages and worse disease-free survival (DFS) in ESCA patients, which can be utilized for prognostic prediction. Furthermore, elevated ANOS1 expression in ESCA correlates with (i) enrichment of epithelial-mesenchymal transition (EMT) based on gene set enrichment analysis (GSEA), (ii) involvement in hypoxia, angiogenesis, WNT signaling pathway, and TGF-β signaling pathway based on gene set variation analysis (GSVA), (iii) presence of the small insertion and deletion (ID) mutational signature ID9, associated with chromothripsis, in single-nucleotide polymorphism (SNP) analysis, (iv) amplification of 11q13.3 based on copy number variants (CNV) analysis, (v) presence of cancer-associated fibroblasts and mesenchymal stromal cells in the tumor microenvironment.

Multi-omic analyses collectively suggest a pivotal role of ANOS1 in accelerating ESCA progression. Subsequent *in vitro* and *in vivo* experiments further support the biological role of ANOS1 in ESCA. Knockdown of ANOS1 significantly impairs the proliferation of ESCA cells, indicating its potential as a novel therapeutic target for ESCA patients. Additionally, we identified potential chemotherapeutics, such as cisplatin, docetaxel, cyclopamine, and vinorelbine, with lower-half maximal inhibitory concentration (IC50) in the high-ANOS1 group.

Materials and methods

Data collection

Clinical information and RNA sequencing data of ESCA were obtained from the TCGA database (https://portal.gdc.cancer.gov/), including 185 cancer samples and 13 noncancerous samples. The gene expression transcripts per million (TPM) RNA sequencing data of normal esophageal tissues (653 cases) from the Genotype-Tissue Expression (GTEx) project were downloaded using UCSC Xena (https://xena.ucsc.edu/). The RNA-sequencing data were normalized into TPM and transformed using a log2 scale. The expression profile of GSE53625, which contains 179 ESCC cases, was obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

Analysis of ANOS1 expression in ESCA and normal tissues

The expression difference of ANOS1 between the tumor and adjacent tissues was analyzed by the two-sample Student's t-test when the data in each group were normally distributed; otherwise, by the Mann-Whitney U test in TCGA ESCA dataset and GSE53625.

Comparison between ANOS1 expression and clinicopathological features

The ANOS1 expression was analyzed according to clinical and pathological features, including age, race, tumor location, histologic type, T stage, N stage, and M stage in the TCGA ESCA dataset.

Clinical utility of ANOS1 in diagnosis

The predictive value of ANOS1 for distinguishing between benign and malignant cases was analyzed using the receiver operating characteristic curve (ROC). The performance of ANOS1 is considered poor when the area under the curve (AUC) ranges from 0.5 to 0.6, fair when it ranges from 0.6 to 0.7, and good when it is greater than 0.7.

Survival analysis

Gene Expression Profiling Interactive Analysis (GEPIA) [8] was used to investigate whether ANOS1 is associated with prognosis in ESCA. Additionally, variables associated with overall survival (OS) were evaluated in the univariable and multivariable Cox regression analyses in the TCGA cohort. Subsequently, a feasible nomogram integrated with ANOS1 was constructed using the R package 'rms' [9]. The discrimination power of the nomogram was assessed using calibration curves.

Differentially expressed gene analysis

Based on the median expression level of ANOS1, patients with ESCA in TCGA were divided into the high- and low-ANOS1 expression groups. The differentially expressed genes (DEGs) between these two groups were filtered

using the thresholds of adjusted *p*-value < 0.05 and |log2-fold-change (FC)| >2, employing the "limma" package in R.

Identification of hub genes

To identify the hub genes among the DEGs, we utilized the Search Tool for the Retrieval of Interacting Genes (STRING) database (https://cn.string-db.org/). We established a protein-protein interaction (PPI) network based on the DEGs using a threshold interaction score of >0.9. The "Degree" algorithm in Cytoscape CytoHubba (v3.9.0) was then used to identify the hub genes within this network. In addition, we employed the GOSemSim package in R for further identification of hub genes [10].

Functional enrichment analysis

Gene ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using Over Representation Analysis (ORA) and Gene Set Enrichment Analysis (GSEA), respectively, with the "clusterProfiler" package in R. Additionally, GSEA was employed to investigate the enrichment of signaling transductions and biological functions using the Msigdb database and HALLMARK gene set through clusterProfiler. Furthermore, Gene Set Variation Analysis (GSVA) was implemented using the HALLMARK gene set and KEGG gene set from the Msigdb database with the "GSVA" package in R [11]. The threshold of statistical significance was set as an adjusted *p*-value < 0.05 and |log2 (FC)| >0.1.

Regulation network

The correlation between the DEGs and a total of 318 transcription factors from http://www.cistrome.org/ was determined using a correlation coefficient (R) < 0.5 and *p*-value = 0.00001. Furthermore, the miRNAs that interact with the DEGs, validated through luciferase reporter assay, were obtained using the R packages "multiMiR" and "mirtarbase" [12]. The long noncoding RNAs (lncRNAs) that interact with these miRNAs were obtained from the "starbase" database [13]. To visualize the correlations among the DEGs, miRNAs, and lncRNAs, we utilized the "ggalluvial" package.

Somatic genomic alternation analysis

The somatic mutation data were downloaded from the TCGA and analyzed using the R package "maftools" [14]. The analysis of somatic genomic alterations with maftools was previously described [15]. For the analysis of single base substitution (SBS) signatures, maftools were utilized. Additionally, the doublet base substitution (DBS) and indel (ID) mutational signatures were estimatjed using the "sigminer" package [16]. The extracted signatures were compared against the catalog of somatic mutations in cancer (COSMIC) mutational signatures (https://cancer.sanger.ac.uk/cosmic). Tumor mutational burden (TMB) values were calculated as the number of mutations per megabase (MB) and categorized as low (TMB < 6), intermediate (6 ≤ TMB < 20), or high (TMB ≥ 20) [17].

Copy number variation analysis

The masked CNV data were downloaded from TCGA, and the common copy number variation (CNV) regions were identified using the GISTIC 2.0 module in GenePattern (https://cloud.genepattern.org/gp/pages/index.jsf).

Tumour microenvironment analysis

The characterization of cellular compositions in the tumor microenvironment of ESCA patients was performed using the "IOBR" package [18]. Subsequently, the correlations between ANOS1 expression and each infiltrating cell type were calculated using the "corrplot" package in R. The association between the infiltrating cell type and OS was evaluated using Cox regression analysis, and the *p*-values for the correlation were transformed into log10 values. Infiltrating cell types with a correlation *p*-value below 0.0001 were chosen and visualized.

The potential therapeutics concerning ANOS1 expression

The R package "pRRophetic" was used to estimate potential chemotherapeutics for treating ESCA patients with high ANOS1 expression [19]. Furthermore, the TIDE website was employed to calculate possible implications for immunotherapy (https://tide.dfci.harvard.edu/).

Cell culture and transfection

In the study, the cell lines utilized included the immortalized human esophageal epithelial cell line (HEEC) and human esophageal squamous cell carcinoma cell lines (TE10, TE13, KYSE150, KYSE450, Eca109), which were obtained from the Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences. These cells were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS, Gibco Company) and 10% penicillin-streptomycin (Sigma-Aldrich). The cells were incubated at 37°C with 5% CO_2.

Plasmid construction

The lentiviral expressing short hairpin RNA (shRNA) targeting three different sequences of ANOS1 gene (shANOS1-1: CTTCCGATCATTATGTCCTAA, shANOS1-2: GCTTCATTCATCGTCCAGGAT, shANOS1-3: CGAGTTCAACTGACTGACATA) and the negative control (TTCTCCGAACGTGTCACGT) were synthesized and cloned into GV811 (pFU-GW-012-SV40-puro), vector with EcoRI sites (purchased from Shanghai Genechem Co., Ltd.), recombinant vector was detected by DNA sequencing.

Lentivirus production and infection

The viral vectors were transfacted into 293T cells using Lipofectamine 2000 (Invitrogen; Thermo Fisher Scientific, Inc.) together with two helper plasmids psPAX2 and pMD2.G. Infectious lentiviruses were harvested 72 h post-transfection, rapid centrifugation to remove cell debris, and then filtered through 0.45 μm cellulose acetate filters. The virus titer was determined by quantitative PCR analysis and was approximately 1×10^9 transducing units (TU)/mL medium. Then, the KYSE150 and KYSE450 cells were infected at a multiplicity of infection of 10 particles per cell. The infected cells were treated with puromycin 72 hours post-infection. After one week of drug selection, stably infected cell lines were obtained.

RNA extraction and quantitative real-time polymerase chain reaction (qRT-PCR)

Total cellular RNA was isolated using TRIzol reagent (Vazyme Biotech, Nanjing, China). The concentration and purity of the extracted RNA were determined using a Bio-Spec-nano spectrophotometer (Shimadzu, Japan). The RNA was then reverse-transcribed into complementary DNA (cDNA) using the Prime Script RT Master Mix reagent (Takara Bio, Dalian, China). Subsequently, the quantitative real-time PCR (qRT-PCR) was carried out using the StepOnePlus real-time PCR system (Thermo Fisher Scientific) and the TB Green® PreMix Ex Taq™ polymerase chain reaction system (Takara Bio, Dalian, China). The $2^{-\Delta\Delta CT}$ method was employed to calculate the expression levels of the target genes. For ANOS1, the primer sequences used were sense: 5'-CCATGACTGGGTTTCAAGTG-3' and antisense: 5'-GGACATAATGATCGGAAGGC-3'. GAPDH was used as a reference gene.

Cell counting kit-8 (CCK-8) assay

Cell proliferation was assessed using a Cell Counting Kit-8 (CCK-8) assay from Beyotime (Shanghai, China) following the provided protocols. Cells (2000 cells per well) were seeded into 96-well plates. At a specific time point, CCK-8 solution was added to each well, and the cells were further incubated at 37°C for 2 hours. The absorbance values were then measured at 450 nm using a microplate spectrophotometer (Thermo, USA). These measurements were used to determine the proliferation capability of ESCC cells.

Clonogenic survival assay

Cells were seeded at a density of 300 cells per well in 6-well plates and cultured in a 5% CO_2 incubator at 37°C for 10 to 14 days. Afterward, proliferating cell colonies were fixed using 1% formaldehyde and stained with 1% crystal violet. The number of colonies containing at least 50 cells was counted, and photographs were taken for documentation purposes.

Xenograft mouse model

Female BALB/c nude mice aged 5-6 weeks were obtained from the Jiangnan University Animal Centre (Wuxi, China). The mice were randomly divided into two groups with three mice in each group: (i) sh-Control and (ii) sh-ANOS1-1. To establish xenografts, the mice were injected with 0.2 mL of PBS containing 5×10^6 cells into the right armpit. On day 14 after injection, the mice were euthanized. Tumor weights were measured upon sacrificing the mice. All animal experiments were approved by

the Ethics Committee of Jiangnan University (JN. No 20230315b0140531 [068]).

Immunohistochemistry analysis

A total of 112 cases of esophageal squamous cell carcinoma (ESCC) and 68 cases of healthy surrounding esophageal tissues, including 48 paired samples of ESCC and adjacent tissues, were gathered for the evaluation of anosmin-1 expression levels utilizing immunohistochemistry (IHC) analysis. The study was conducted by the approved protocol of the Ethics Committee of Shanghai Outdo Biotech Company (JS W-03-01). Written informed consent was obtained from all participants, and the study adhered to the principles outlined in the Helsinki Declaration.

Specific antibodies against KAL1 (NBP3-03937, Novus) were used as the primary antibody at a dilution of 1:500, with a Rabbit monoclonal [EPR20115] antibody serving as the secondary antibody in the IHC assays. The assays were performed following the manufacturer's instructions using a biotin assay system (Beijing Zhongshan Jinqiao, China, Cat. Nos. PV-9001, PV-9002).

For the IHC analysis, a grading system was employed. An IHC score of 0 was assigned for the absence of staining, a score of 1 for yellow staining, and a score of 2 for brown staining. Based on the percentage of positively stained tumor cells in the visual field, a score of 0 was given for less than 1% of cells, a score of 1 for 1-25%, a score of 2 for 25-75%, and a score of 3 for 75-100%. The overall score was calculated by multiplying the intensity score by the percentage of positive cells.

Statistical analysis

Statistical analyses were conducted using R (4.1.3). Continuous variables were compared using the Student's t-test (for normally distributed data) or a Mann-Whitney U test (for non-normally distributed data). Paired samples were compared using the paired sample t-test when the data met the assumptions of normality and homogeneity of variances, otherwise using the Wilcoxon test. Alternatively, categorical data were analyzed using either the chi-square test or Fisher's exact test, based on the specific demands of the test. The correlation between continuous variables is evaluated through the Pearson correlation test and simple linear regression analysis. Kaplan-Meier survival analysis and the log-rank test were performed to calculate the *p*-value. A significance level of $P < 0.05$ was considered statistically significant unless otherwise specified.

Results

Analysis of ANOS1 expression in ESCA and normal tissues

The expression of ANOS1 was significantly higher in the TCGA ESCA dataset compared to normal esophageal tissues in GTEx ($P < 0.001$, **Figure 1A**). A similar result was observed in GSE53625, where ANOS1 expression was upregulated in tumors ($P < 0.001$, **Figure 1B**). In normal esophageal tissue, the expression of ANOS1 was lower in the mucosa compared to the muscularis ($P < 0.001$, **Figure 1C**).

Furthermore, the associations between ANOS1 expression and clinicopathologic features were analyzed. The clinical characteristics of the patients concerning ANOS1 expression in the TCGA ESCA dataset are shown in Table S1. There was no significant difference in ANOS1 expression based on age (**Figure 1D**). ANOS1 levels were higher in Asian patients than in White patients ($P < 0.001$, **Figure 1E**). When the tumor was located in the lower third of the esophagus, ANOS1 expression was lower compared to the middle third of the esophagus ($P < 0.001$, **Figure 1F**). ANOS1 expression was significantly increased in squamous cell carcinoma compared to adenocarcinoma ($P < 0.001$, **Figure 1G**), and this result was also observed in Whites ($P < 0.001$, Figure S1A). Compared to the T0/T1 stage, ANOS1 expression was increased in the T2, T3, and T4 stages ($P < 0.001$, $P < 0.001$, $P = 0.006$, respectively, **Figure 1H**). Similar results were observed in Whites, with increased ANOS1 expression in the T2, T3, and T4 stages ($P < 0.001$, Figure S1B). No significant difference in ANOS1 expression was observed according to the N stage and M stage (**Figure 1I** and **1J**). There was a trend that ANOS1 expression was enhanced in the N2 stage compared to the N1 stage in squamous cell carcinoma ($P = 0.087$, Figure S1C). In adenocarcinoma, ANOS1 expression was increased in the T2 and T3 stages

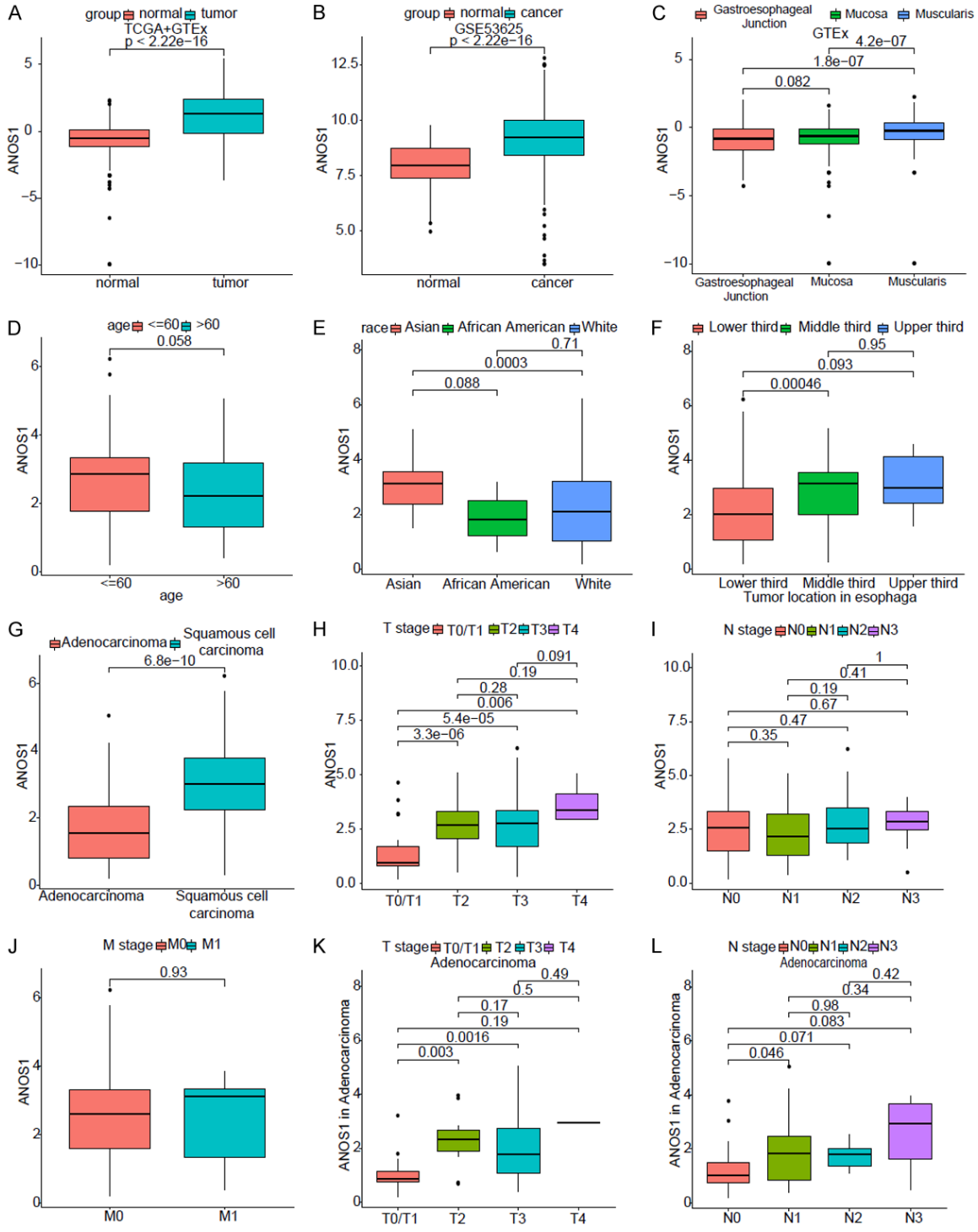


Figure 1. The association of ANOS1 with clinicopathological parameters in ESCA patients. The comparison of ANOS1 expression between ESCA tissue and normal tissue in TCGA and GTEx datasets (A) and GSE53625 dataset (B). The expression of ANOS1 according to the anatomical sites of the esophagus in GTEx (C). The expression of ANOS1 based on age at diagnosis (D), race (E), tumor location (F), histologic type (G), T stage (H), N stage (I), and M stage (J) in the TCGA ESCA dataset. The expression of ANOS1 according to the T stage (K) and N stage (L) in esophageal adenocarcinoma.

compared to the T0/T1 stage (P = 0.003, P = 0.002, respectively, **Figure 1K**), and in the N1 stage compared to the N0 stage (P = 0.046, **Figure 1L**).

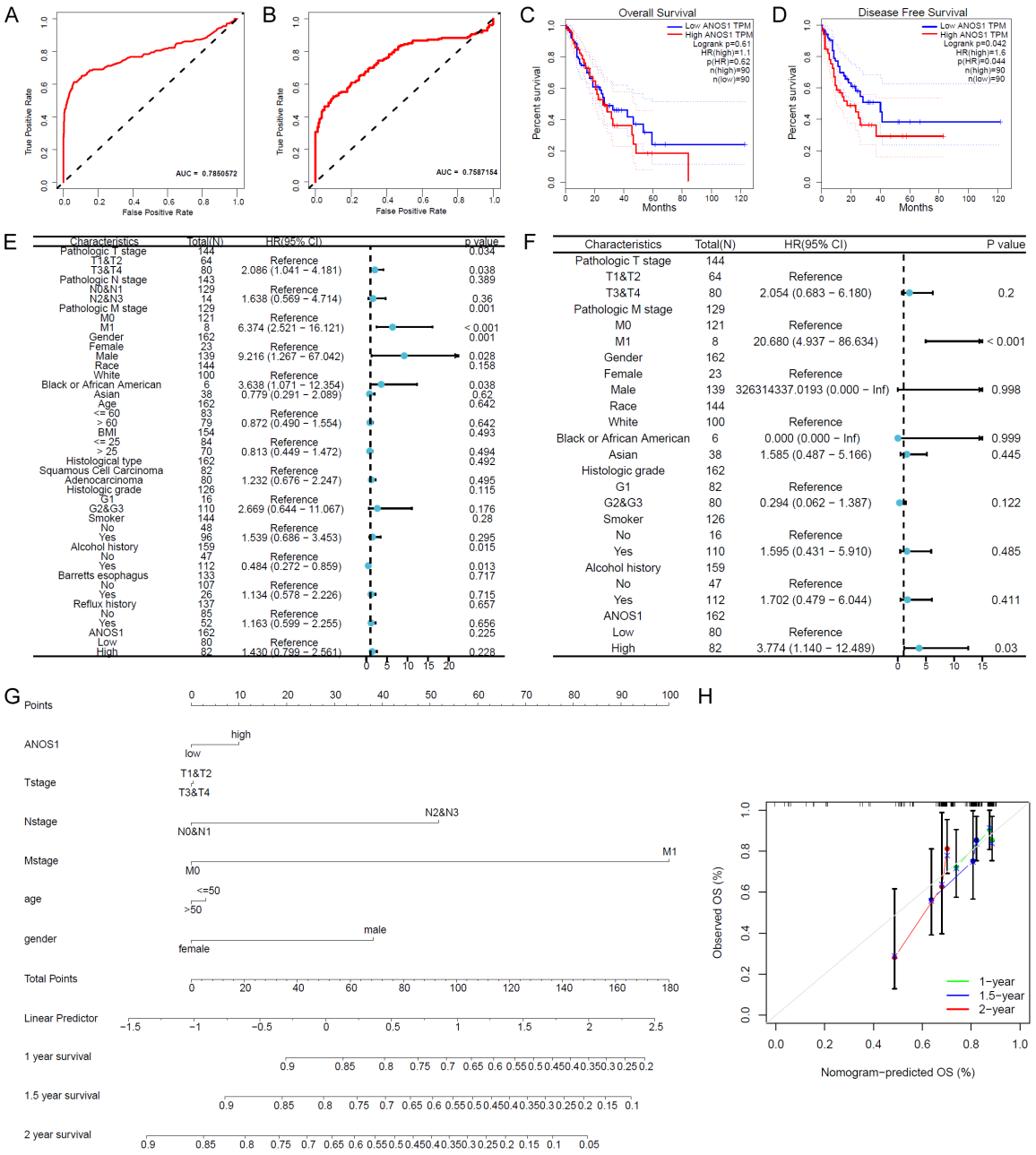


Figure 2. The clinical utility of ANOS1 in diagnosis and prognosis. The predictive value of ANOS1 for ESCA diagnosis was analyzed using ROC in TCGA and GTEx (A) and GSE53625 (B). Expression of ANOS1 in ESCA patients and association with overall survival (C) and disease-free survival (D). Variables associated with overall survival were evaluated in the univariable (E) and multivariable (F) Cox regression analyses in TCGA cohort. (G) A nomogram constructed to predict overall survival rates at 1, 1.5 and 2 years. (H) The nomogram calibration curves on consistency between predicted and observed 1-, 1.5-, and 2-year overall survival. Low and high ANOS1 expression was defined according to the median mRNA levels of respective cohort.

Clinical utility of ANOS1 in diagnosis and prognosis

Since the expression of ANOS1 was significantly enhanced in ESCA, its diagnostic efficacy for ESCA was evaluated using ROC curves. The performance of ANOS1 expression in distinguishing between normal tissue and tumor was found to be good in the TCGA and GTEx datasets (AUC = 0.79, **Figure 2A**) as well as in the GSE53625 dataset (AUC = 0.76, **Figure 2B**). Although ANOS1 expression did not show a

correlation with overall survival (OS) of patients (**Figure 2C**), it was significantly associated with disease-free survival (DFS) (P = 0.04, **Figure 2D**) in the TCGA ESCA dataset. Additionally, ANOS1 expression was significantly associated with worse OS in Asians (Figure S1D).

Furthermore, univariate analysis revealed that pathologic T stage, pathologic M stage, gender, and alcohol history were significantly related to OS (**Figure 2E**). Multivariate Cox regression analysis demonstrated that ANOS1 expression and metastasis remained statistically significant predictors (P < 0.05, **Figure 2F**). A nomogram was constructed to predict 1-, 1.5-, and 2-year overall survival in ESCA (**Figure 2G**). The nomogram calibration plot (**Figure 2H**) shows that the nomogram was well calibrated, indicating that the predicted probabilities for each year closely matched the observed probabilities.

Identification of DEGs and hub genes

To further investigate the mechanism of ANOS1 in cancer progression in ESCA, DEGs were identified between the high-ANOS1 and low-ANOS1 groups. Differential expression analysis revealed 130 DEGs, with 61 genes upregulated and 69 genes downregulated in the high-ANOS1 group (Figures S2 and **3A**). The top 10 upregulated and downregulated genes are highlighted in **Figure 3A**. Furthermore, to identify hub genes among the DEGs, the PPI network was analyzed using the "Degree" algorithm in Cytoscape CytoHubba (**Figure 3B**). Additionally, a "GOSemSim" analysis was conducted (**Figure 3C**). And the top 5 hub genes identified were TFAP2C, NEUROG3, SOX2, ZNF750, and GAS1. Then, the intersection of hub genes derived from both "GOSemSim" and "Degree" analyses, including SOX2, WNT3A, and APOBEC1, were listed in Table S2. To explore the potential impact of copy number variations (CNVs) on the expression of these hub genes, the CNV status of the hub genes was assessed (**Figure 3D**). It was found that some hub genes, such as SOX2 and WNT3A, may be upregulated due to amplification. Interestingly, the expression of ANOS1 showed a strong negative association with the expression of Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 1 (APOBEC1) (**Figure 3E**). The APOBEC family is responsible for multiple cytosine deaminations in the genome [20]. Moreover, the expression of ANOS1 was positively associated with the expression of WNT3A (R = 0.47, P < 0.0001, Figure S3A) and WNT7A (R = 0.36, P < 0.0001, Figure S3B).

Functional enrichment analysis

Then, GO and KEGG analyses were performed for the DEGs. The GO analysis revealed that the DEGs were involved in the development and morphogenesis of the embryo, as shown by ORA (Figure S4A) and GSEA (Figure S4B). Additionally, the GSEA results for GO analysis demonstrated activated transcription activity along with suppressed metabolism processes and immune responses (Figure S4B). In the KEGG analysis, the estrogen signaling pathway was found to be significantly enriched according to ORA (Figure S4C), while metabolic pathways were significantly suppressed based on GSEA (Figure S4D and S4E). Furthermore, the GSEA analysis using MsigDB showed a significantly enhanced WNT signaling transduction (Figure S4F). The HALLMARK analysis revealed a significantly increased EMT (Figure S4G). Moreover, GSVA identified an overrepresentation of EMT, WNT signaling pathway, TGFβ signaling pathway, and other pathways (**Figure 4A** and **4B**). GSVA of the KEGG Pathways further uncovered suppression of metabolic pathways in the high-ANOS1 groups (**Figure 4B**).

Regulation network

Since the GO analysis showed activation of transcription activity, the regulation of transcription was further explored, resulting in the identification of 21 transcription factors based on the DEGs (**Figure 5A**). Additionally, the expressions of SOX2, TCF7L1, and TFAP2C were strongly positively associated with ANOS1 expression (**Figure 5B-D**). Conversely, PDX1 levels showed a strong negative association with ANOS1 levels (**Figure 5E**). Moreover, post-transcriptional regulations by miRNA and lncRNA were inferred (**Figure 5F**), revealing the involvement of LINC00641 and TUG1 in the post-transcriptional regulation of the DEGs. Notably, the expression of ANOS1 was found to be positively associated with the expression of LINC00641 (R = 0.13, P = 0.02, Figure S3C) and TUG1 (R = 0.45, P < 0.01, Figure S3D).

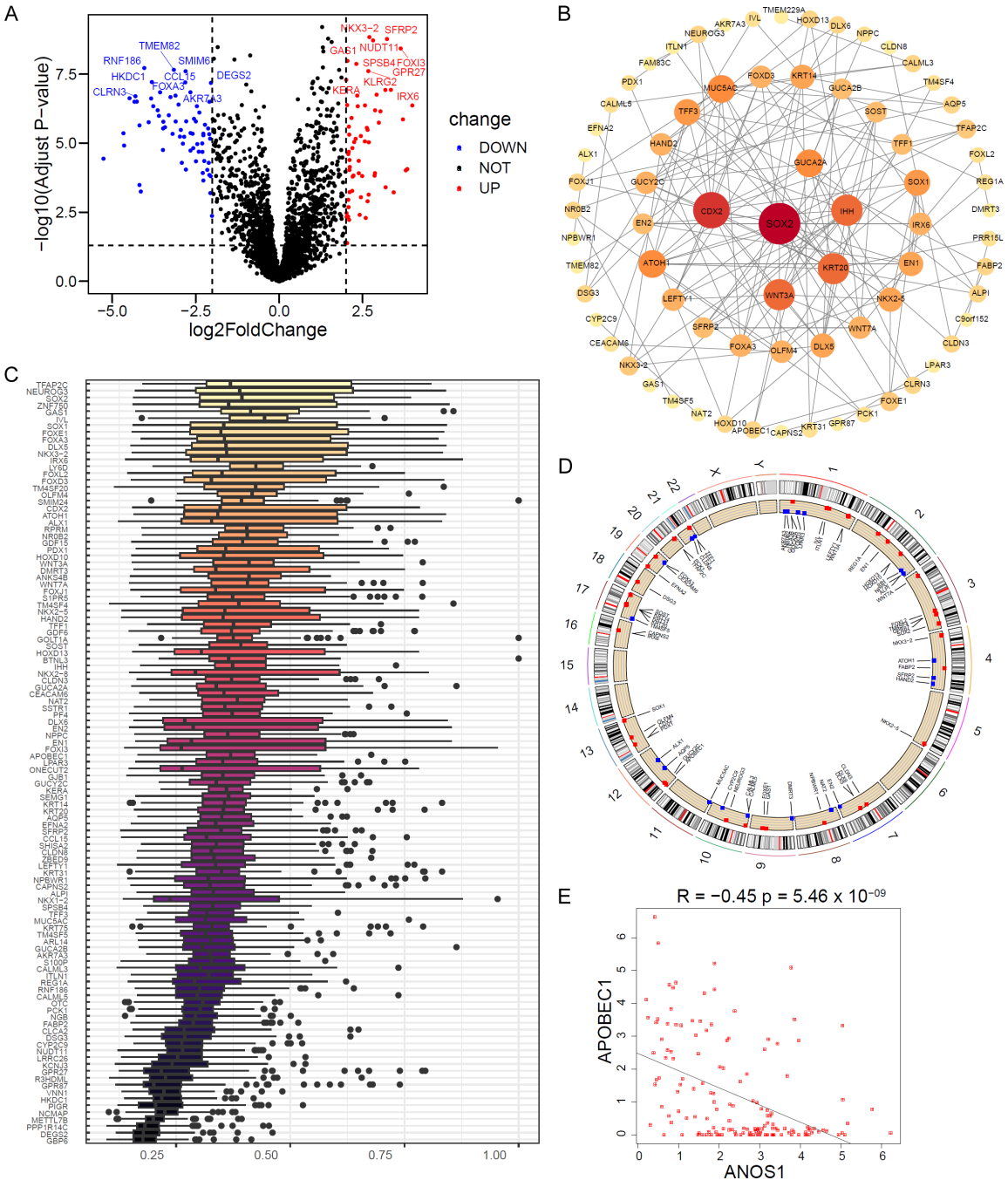


Figure 3. Identification of differentially expressed genes and hub genes. (A) The volcano plot exhibiting 130 differentially expressed genes (DEGs), with the top 10 upregulated genes highlighted in red and the top 10 downregulated genes highlighted in blue. The hub genes among DEGs identified by "Degree" in Cytoscape CytoHubba (B) and "GOSemSim" analysis (C). (D) The CNV profiles of hub genes are displayed as amplification (segment mean >0.2) in red and deletion (segment mean < -0.2) in blue. (E) The association between the expression of ANOS1 and APOBEC1 and the Pearson's coefficient Cor r as and *p*-value are represented.

Somatic mutation analysis according to the ANOS1 expression

In the somatic mutation landscape of the low-ANOS1 group, the top three variant classifications were missense mutation, frameshift deletion, and nonsense mutation. SNP was the most frequent variant type, with single nucleotide variants (SNVs) predominantly occurring as C>T mutations (**Figure 6A**). Similarly, in the

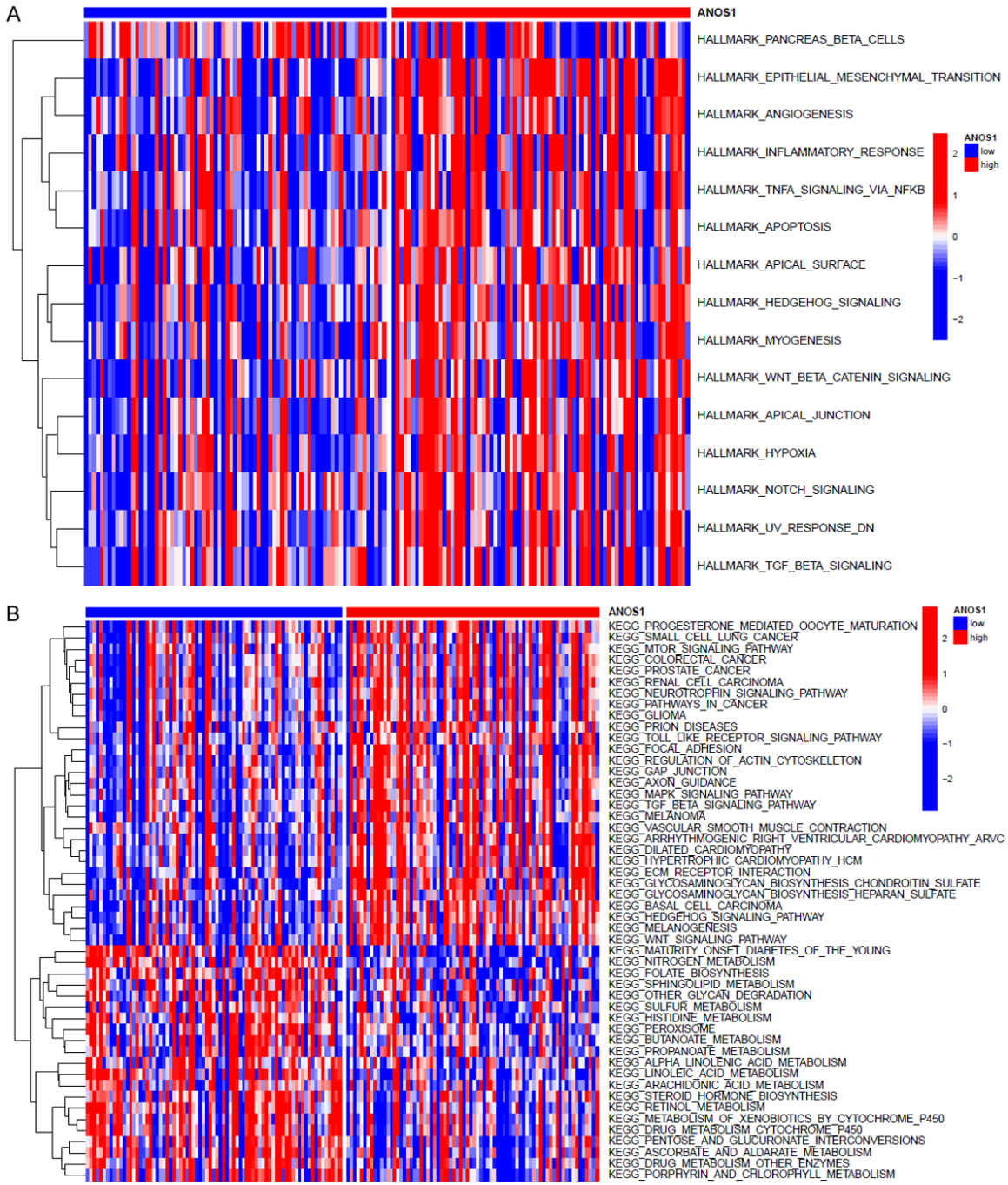


Figure 4. GSVA based on ANOS1 groups. HALLMARK (A) and KEGG (B) pathways according to the ANOS1 expression groups by GSVA.

high-ANOS1 group, the top three variant classifications were missense mutation, nonsense mutation, and frameshift deletion. SNP was also the most frequent variant type, with SNVs of C>T occurring predominantly (**Figure 6B**). The median number of variants per sample was 98 in the low-ANOS1 group and 81.5 in the high-ANOS1 group (**Figure 6A** and **6B**). Consistent with the variant count, the tumor mutational burden (TMB) was calculated as 1.96 mutations per megabase (Mb) in the low-ANOS1 group and 1.63 mutations per Mb in the high-ANOS1 group (Figure S5C and S5D). The top five mutated genes in the low-ANOS1 group

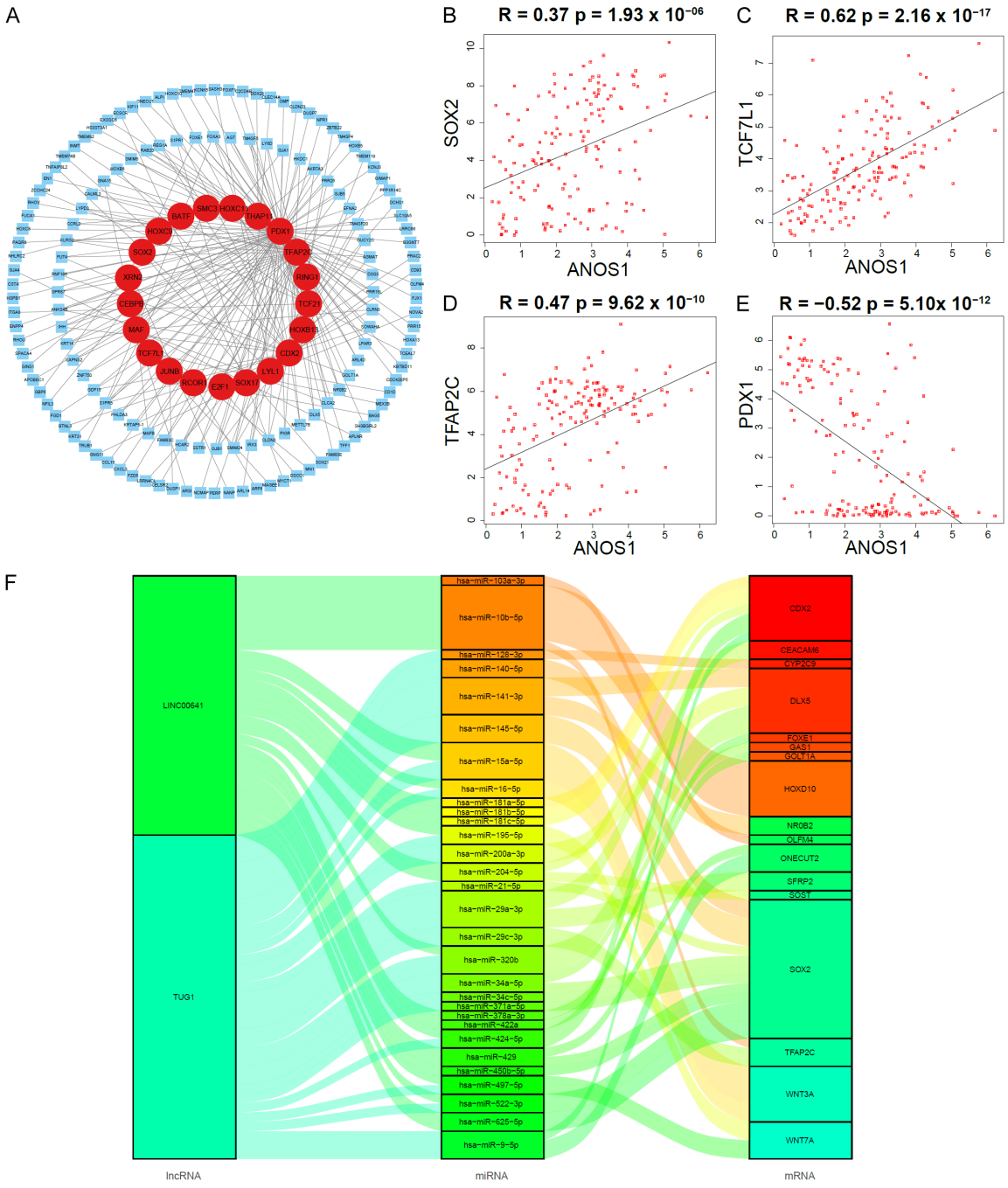
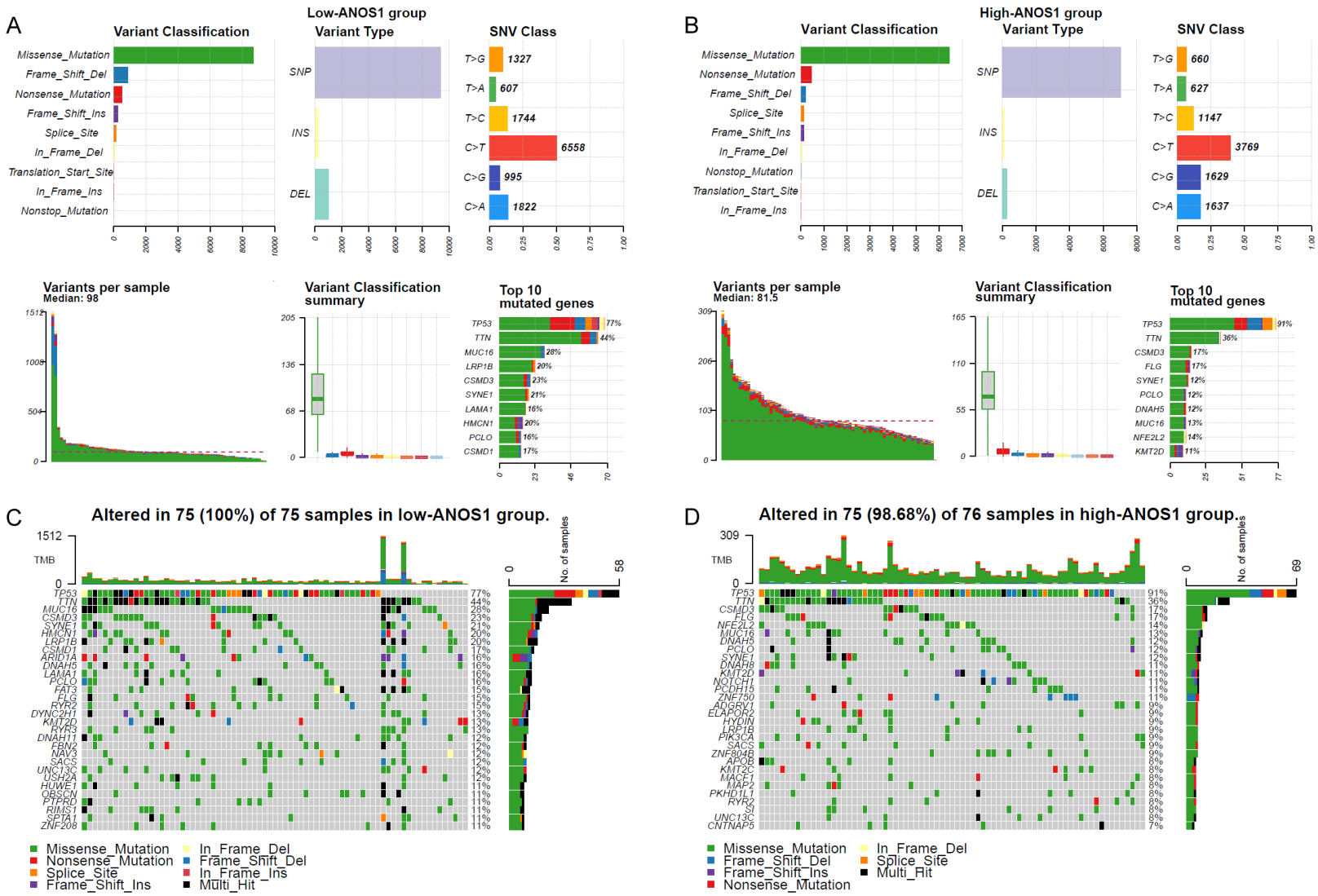


Figure 5. The regulation network of DEGs. (A) The transcriptional regulation network among DEGs, transcription factors in red; other genes in blue. The correlation between ANOS1 and differential expressed transcription factors including SOX2 (B), TCF7L1 (C), TFAP2C (D), and PDX1 (E). (F) The post-transcriptional regulation network of DEGs.

were TP53, TTN, MUC16, CSMD3, and SYNE1, while in the high-ANOS1 group, they were TP53, TTN, CSMD3, FLG, and NFE2L2 (**Figure 6C** and **6D**). All 75 patient samples (100%) in the low-ANOS1 group had somatic mutations (**Figure 6C**), and 75 samples (98.68%) in the high-ANOS1 group also had somatic mutations (**Figure 6D**).

Since cancer progression is driven by proliferatively advantageous driver mutations rather than selectively neutral passenger mutations


A
Low-ANOS1 group
Variant Classification
Variant Type
SNV Class
Missense_Mutation
Frame_Shift_Del
Nonsense_Mutation
Frame_Shift_Ins
Splice_Site
In_Frame_Del
Translation_Start_Site
In_Frame_Ins
Nonstop_Mutation
SNP
INS
DEL
T>G 1327
T>A 607
T>C 1744
C>T 6558
C>G 995
C>A 1822
Variants per sample
Median: 98
Variant Classification summary
Top 10 mutated genes
TP53 77%
TTN 44%
MUC16 28%
LRP1B 20%
CSMD3 23%
SYNE1 21%
LAMA1 16%
HMCN1 20%
PCLO 16%
CSMD1 17%
B
High-ANOS1 group
Variant Classification
Variant Type
SNV Class
Missense_Mutation
Nonsense_Mutation
Frame_Shift_Del
Splice_Site
Frame_Shift_Ins
In_Frame_Del
Nonstop_Mutation
Translation_Start_Site
In_Frame_Ins
SNP
INS
DEL
T>G 660
T>A 627
T>C 1147
C>T 3769
C>G 1629
C>A 1637
Variants per sample
Median: 81.5
Variant Classification summary
Top 10 mutated genes
TP53 91%
TTN 36%
CSMD3 17%
FLG 17%
SYNE1 12%
PCLO 12%
DNAH5 12%
MUC16 13%
NFE2L2 14%
KMT2D 11%
C
Altered in 75 (100%) of 75 samples in low-ANOS1 group.
TMB
No. of samples
Missense_Mutation
Nonsense_Mutation
Splice_Site
Frame_Shift_Ins
In_Frame_Del
Frame_Shift_Del
In_Frame_Ins
Multi_Hit
D
Altered in 75 (98.68%) of 76 samples in high-ANOS1 group.
TMB
No. of samples
Missense_Mutation
Frame_Shift_Del
Frame_Shift_Ins
Nonsense_Mutation
In_Frame_Del
Splice_Site
Multi_Hit

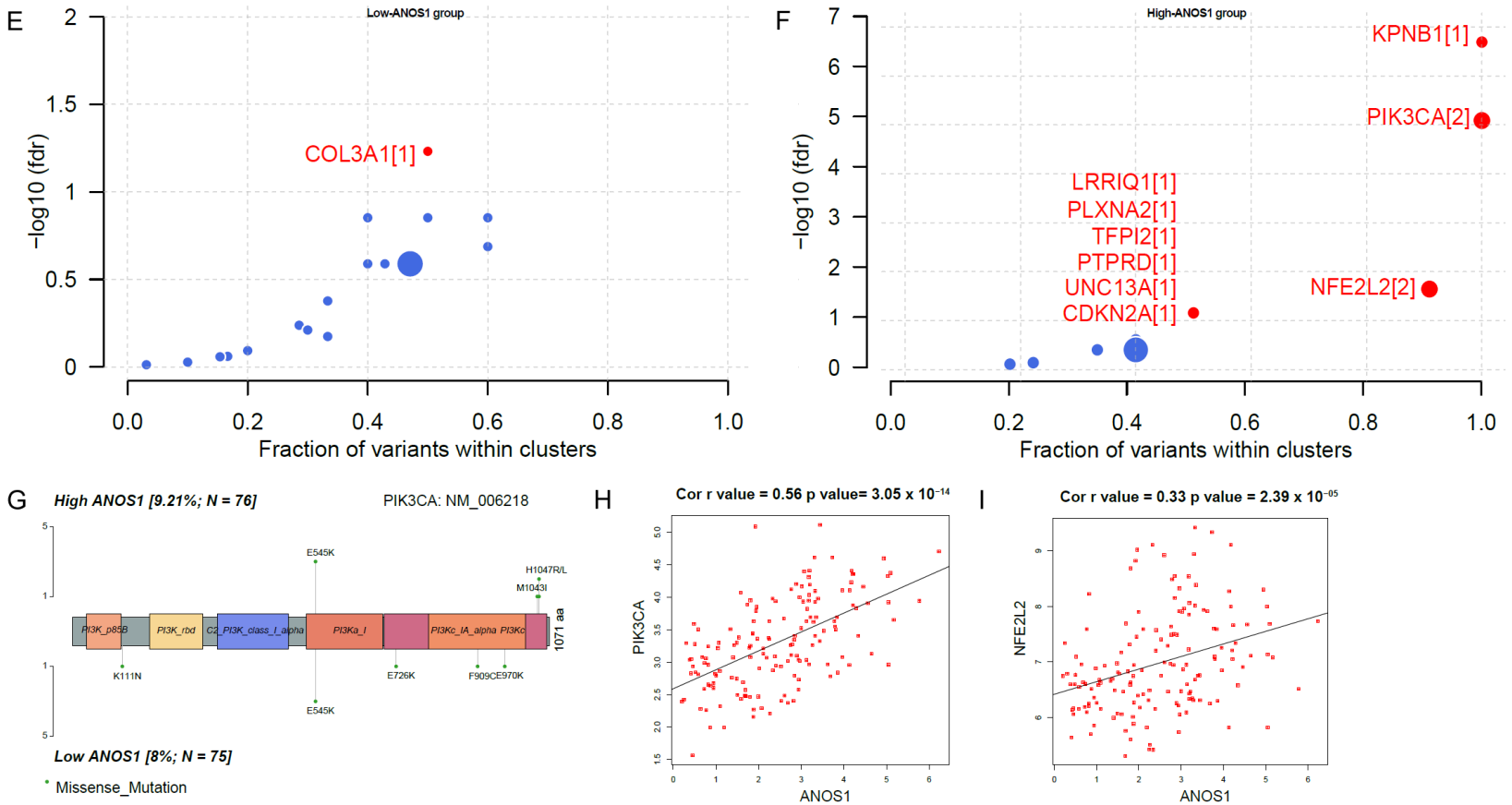


Figure 6. Somatic mutation analysis according to the ANOS1 expression. The summary of the somatic mutation, including variant classification, variant type and SNV class, mutation load for each sample, and variant classification type in the low-ANOS1 group (A) and the high-ANOS1 group (B). The oncoplot displays the somatic mutation landscape of top 20 genes in the low-ANOS1 group (C) and the high-ANOS1 group (D). Driver genes identified by oncodrive function in maftools in the low-ANOS1 group (E) and in the high-ANOS1 group (F). (G) Lollipop plot displays mutation distribution and protein domains for PI3KCA in high-/low-AONS1 groups with labelled recurrent hotspots. The correlation between the expression of ANOS1 and driver genes including PIK3CA (H) and NFE2L2 (I). SNVs, single-nucleotide variants.

[21, 22], the driver mutations were assessed (**Figure 6E** and **6F**). In the low-ANOS1 group, the driver gene COL3A1 was identified, while in the high-ANOS1 group, the top driver genes were KPNB1, PIK3CA, and NFE2L2 (**Figure 6E** and **6F**). Additionally, CDKN2A was also identified as a mutation driver in the high-ANOS1 group, which is consistent with previous reports [23]. Notably, the H1047R, H1047L, and M1043I mutations of PIK3CA were exclusively identified in the high-ANOS1 group (**Figure 6G**). Interestingly, the expression of ANOS1 showed a strong positive correlation with the expression of PIK3CA (**Figure 6H**) and a positive association with the expression of NFE2L2 (**Figure 6I**). Furthermore, the top five pathways affected by mutations were the RTK-RAS, WNT, NOTCH, Hippo, and PI3K signaling pathways (Figure S5A and S5B). The co-occurrence and mutual exclusivity of mutations were displayed in Figure S5E and S5F.

Mutational signature analysis

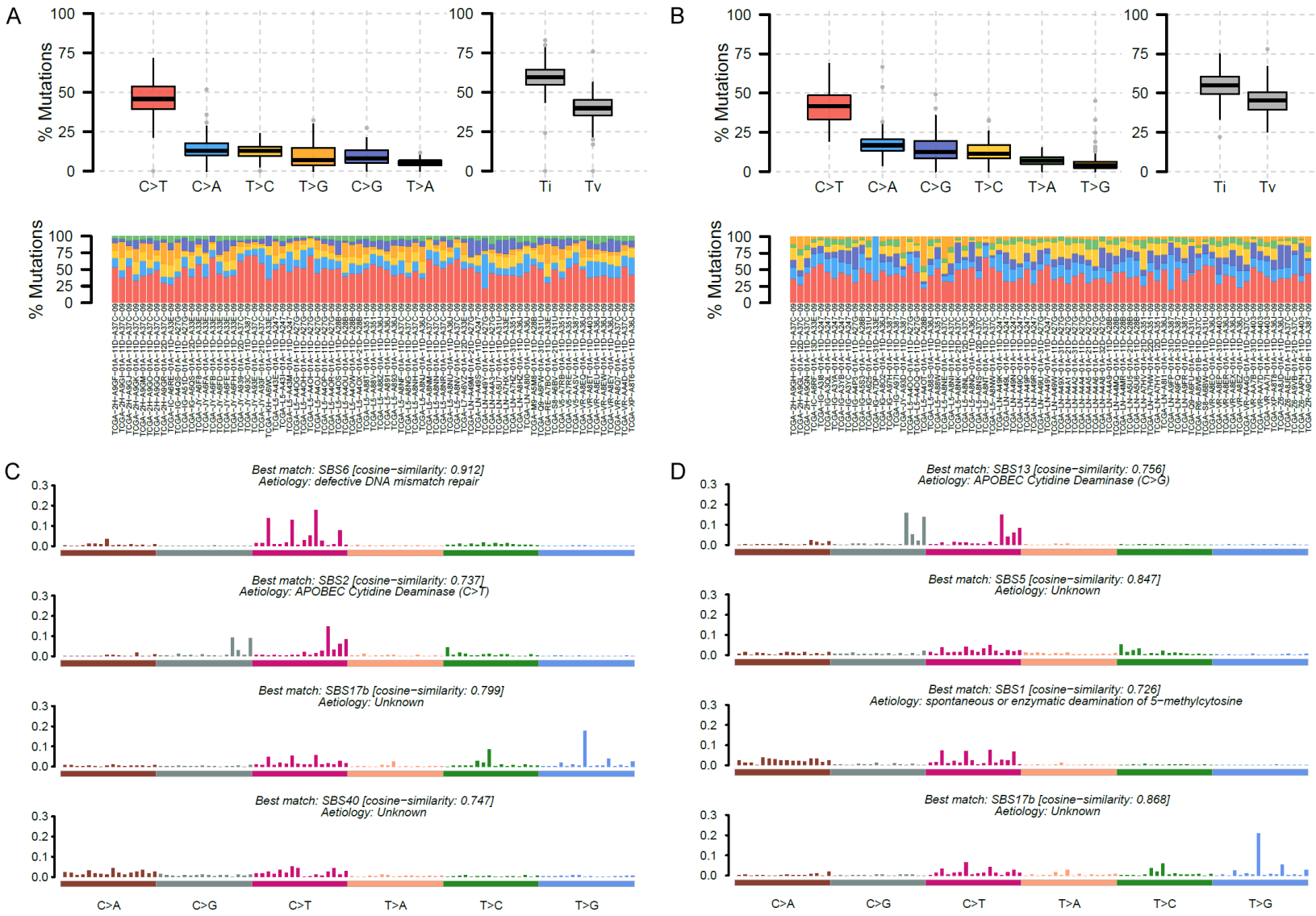
SNPs are divided into two categories based on base substitutions, including two types of transitions (Ti) (A>G/G>A and T>C/C>T) and four types of transversions (Tv) (C>A/A>C, C>G/G>C, T>A/A>T, and T>G/G>T). **Figure 7A** and **7B** show the fractions of these conversions in each group. The top three conversions in the low-ANOS1 group are C>T, C>A, and T>C, while in the high-ANOS1 group, they are C>T, C>A, and C>G (**Figure 7A** and **7B**). Given that the frequency of each SNV class differs between the groups, the single base substitution (SBS) signatures contributing to the high mutation rate were further identified using maftools. In the low-ANOS1 group, SBS6, SBS2, SBS17b, and SBS40 signatures were identified (**Figure 7C**), whereas in the high-ANOS1 group, the signatures of SBS13, SBS5, SBS1, and SBS17b were observed based on the SBS signature analysis (**Figure 7D**) [24]. The SBS2 signature is predominantly characterized by C>T mutations, while the SBS13 signature is characterized by C>G and C>A mutations [25], which partially reflects the higher frequency of C>G mutations in the high-ANOS1 group compared to the low-ANOS1 group (**Figure 7D**). The APOBEC signatures (SBS2 in the low-ANOS1 group and SBS13 in the high-ANOS1 group) were identified, indicating the presence of APOBEC-associated mutations. To further explore the status of APOBEC-associated mutations in the low/high-ANOS1 groups of ESCA patients, the analysis was performed (**Figure 7E** and **7F**). In the low-ANOS1 group, 12% of patients (9 out of 75 samples) showed enrichment for APOBEC-associated mutations. However, there was no difference in the mutation load between APOBEC-enriched and non-APOBEC-enriched samples (**Figure 7E**). In contrast, in the high-ANOS1 group, 30% of patients (23 out of 76 samples) exhibited enrichment for APOBEC-associated mutations, and the mutation load was significantly higher in the APOBEC-enriched samples compared to the non-APOBEC-enriched samples (**Figure 7F**). It is worth noting that APOBEC enzymes deaminate cytidines and predominantly target a tCw motif [26]. The tCw loads in the APOBEC-enriched samples were 0.26 in the low-ANOS1 group and 0.31 in the high-ANOS1 group (**Figure 7E** and **7F**).

Moreover, the doublet base substitution (DBS) mutational signatures were extracted in the low-ANOS1 group (**Figure 8A**) and the high-ANOS1 group (**Figure 8B**). These extracted DBS signatures were then compared against the COSMIC DBS signatures, and the respective cosine similarity and etiology of each DBS signature were summarized in Table S3. In the low-ANOS1 group, the presence of DBS4 (similarity 0.703) and DBS2 (similarity 0.966) were identified (**Figure 8A**). In the high-ANOS1 group, DBS1 (similarity 0.666), DBS2 (similarity 0.626), and DBS4 (similarity 0.642) were identified (**Figure 8B**).

Additionally, the ID mutational signatures were explored based on the expression level of ANOS1 (**Figure 8C** and **8D**). The extracted ID signatures were compared against the COSMIC ID signatures, and their respective cosine similarity values were summarized in Table S4. In the low-ANOS1 group, the presence of ID2 was identified with the highest similarity (0.809) (**Figure 8C**). It has been reported that ID2 is elevated in tumors with defective DNA mismatch repair [27]. In the high-*ANOS1* group, ID8 (similarity of 0.742) and ID9 (similarity of 0.758) were identified (**Figure 8D**).

Copy number variant analysis

GISTIC analysis revealed significant regions of recurrent focal chromosomal copy number loss/deletion and gain/amplification. In the


A
B
% Mutations
C>T
C>A
T>C
T>G
C>G
T>A
Ti
Tv
C
D
Best match: SBS6 [cosine−similarity: 0.912]
Aetiology: defective DNA mismatch repair
Best match: SBS2 [cosine−similarity: 0.737]
Aetiology: APOBEC Cytidine Deaminase (C>T)
Best match: SBS17b [cosine−similarity: 0.799]
Aetiology: Unknown
Best match: SBS40 [cosine−similarity: 0.747]
Aetiology: Unknown
Best match: SBS13 [cosine−similarity: 0.756]
Aetiology: APOBEC Cytidine Deaminase (C>G)
Best match: SBS5 [cosine−similarity: 0.847]
Aetiology: Unknown
Best match: SBS1 [cosine−similarity: 0.726]
Aetiology: spontaneous or enzymatic deamination of 5−methylcytosine
Best match: SBS17b [cosine−similarity: 0.868]
Aetiology: Unknown

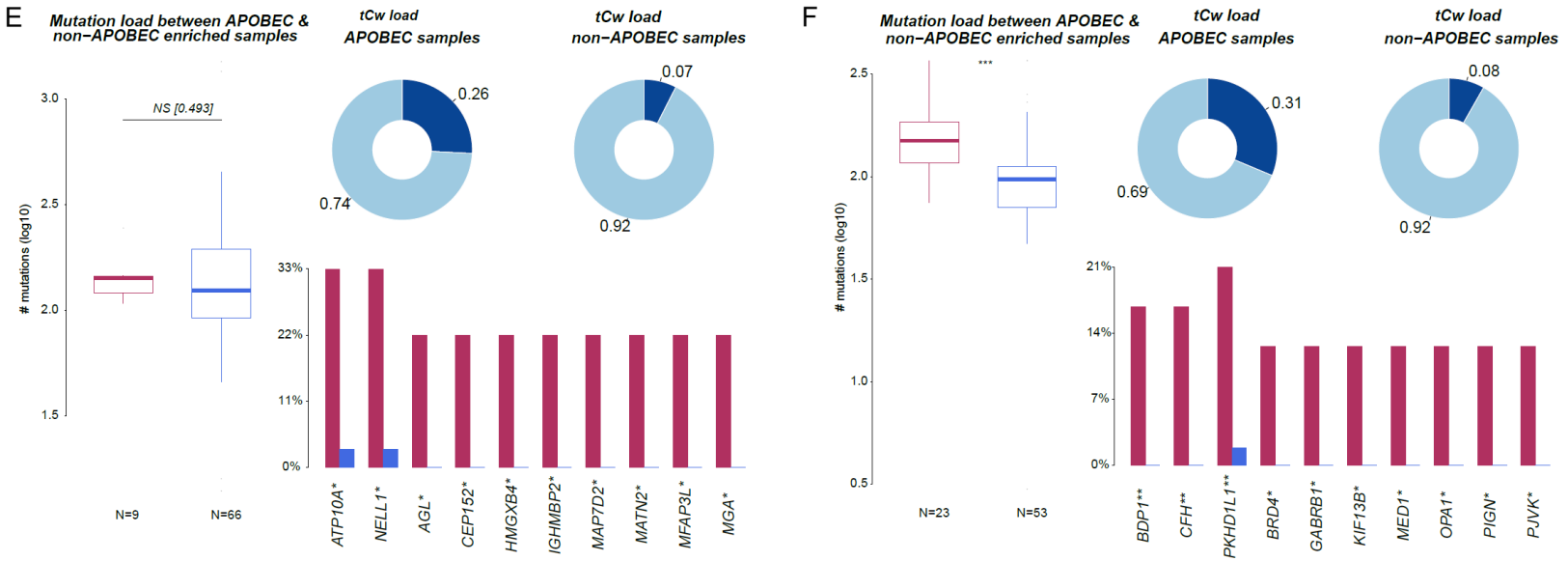


Figure 7. Mutational signature analysis. The distribution of six conversions of transition and transversion is plotted in the low-ANOS1 group (A) and in the high-ANOS1 group (B). Single base substitution (SBS) signatures matched against the validated COSMIC v3 signatures in the low-ANOS1 group (C) and the high-ANOS1 group (D). The APOBEC enrichment analysis in the low-ANOS1 group (E) and the high-ANOS1 group (F). The differences in mutation load between APOBEC-enriched and non-enriched samples are displayed in the boxplots in the left panel. The standard deviation is demonstrated by the error bars. The top 10 differentially mutated genes between APOBEC-enriched and non-enriched samples are displayed in the bar plots. SNVs, single-nucleotide variants.

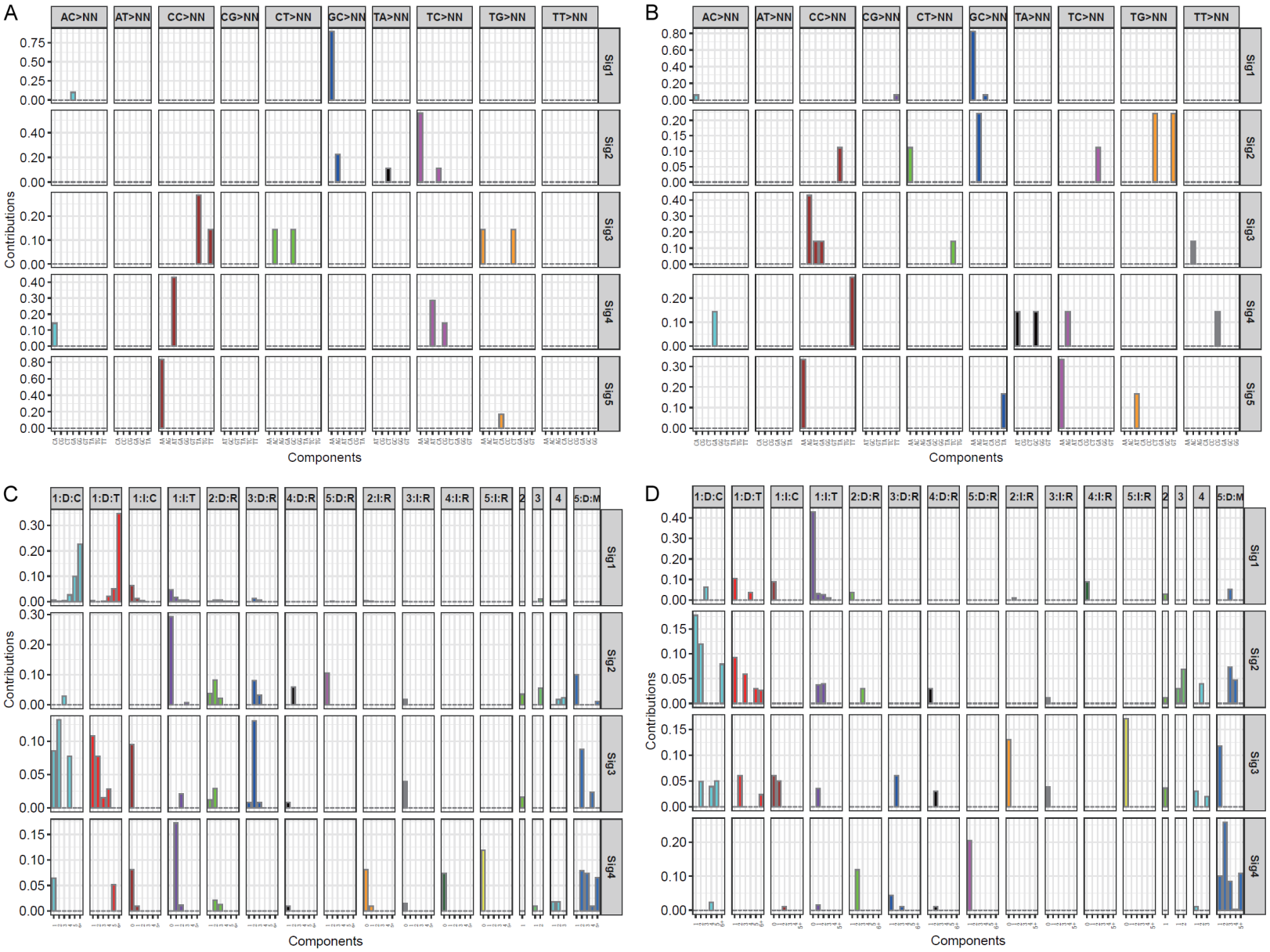


Figure 8. Doublet base substitution and small insertion and deletion mutational signature analysis. Doublet base substitution (DBS) signatures in the low-ANOS1 group (A) and the high-ANOS1 group (B). The small insertion and deletion (ID) mutational signatures in the low-ANOS1 group (C) and the high-ANOS1 group (D).

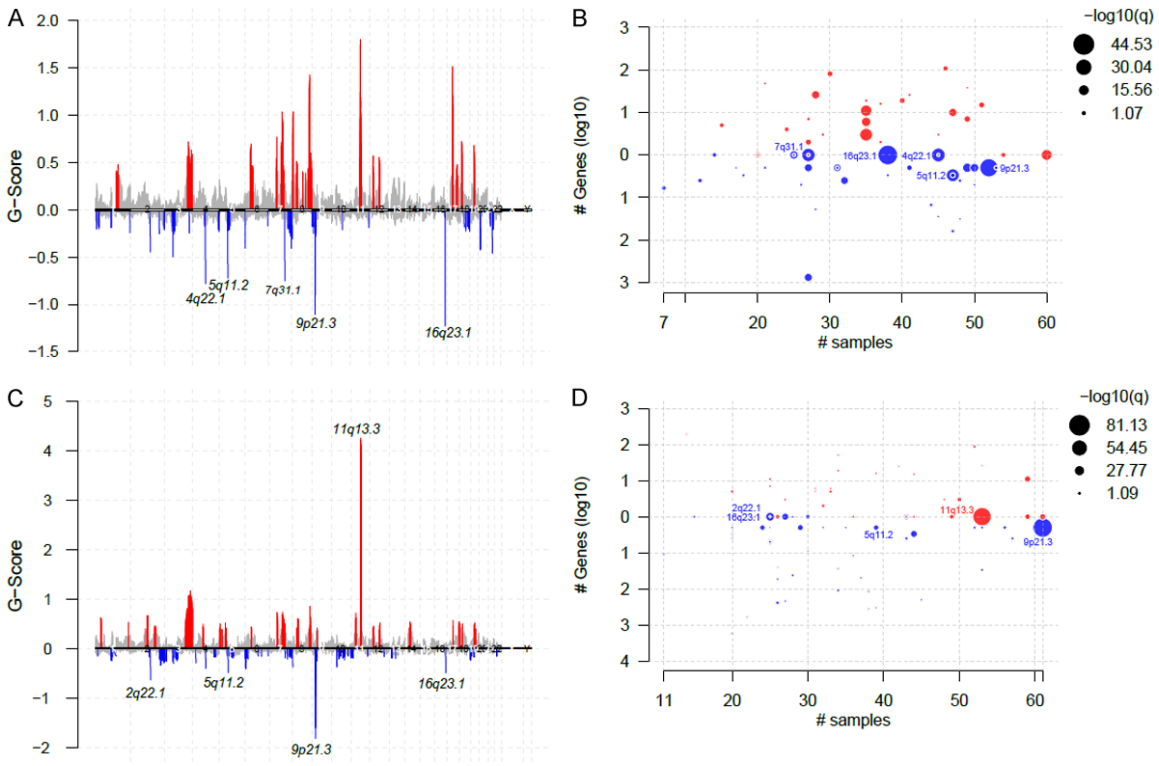


Figure 9. Copy number variant analysis. G-scores of significant recurrent copy number variant (CNV) assigned by GISTIC are depicted along the chromosome in the low-ANOS1 group (A) and in the high-ANOS1 group (C). GISTIC results plotted as recurrent CNV, mutated samples, and genes involved within the recurrent CNV in the low-ANOS1 group (B) and in the high-ANOS1 group (D). The bubble sizes are proportional to -log10 (Q-value).

low-ANOS1 group, several deletions were identified, including 9q21.3 (CDKN2A), 16q23.1 (WWOX), 4q22.1 (CCSER1), and 5q11.2 (PDE4D, PART1) (**Figures 9A**, **9B**, S6A and S6B). In the high-ANOS1 group, the most prevalent amplification was observed in 11q13.3, while the most prevalent deletion corresponded to 9q21.3 (CDKN2A) (**Figures 9C**, **9D**, S6C and S6D). The loss of CDKN2A in both groups is consistent with previous studies [28]. Additionally, the expression of CDKN2A was found to be negatively associated with the expression of ANOS1 (R = -0.23, P = 0.04, Figure S3E). Furthermore, it is worth noting that the 11q13.3 region harbors genes such as FGF3/4/19, CCND1, and FADD, among others [29]. Intriguingly, the expression of *ANOS1* was positively associated with the expression of *FGF3* (R = 0.27, P = 0.001, Figure S3F), *FGF4* (R = 0.25, P = 0.002, Figure S3G), *FGF19* (R = 0.22, P = 0.01, Figure S3H), *CCND1* (R = 0.26, P = 0.001, Figure S3I), *FADD* (R = 0.26, P = 0.001, Figure S3J).

Tumour microenvironment analysis based on the ANOS1 expression

First, the correlation between cell types or scores and OS was calculated using univariate Cox regression analysis. Eosinophils_CIBERSORT and T cells γδ (Tgd)_cells_xCell showed a positive association with prognosis [HR (95% CI) 1.85 (1.16-2.95), P = 0.01; HR (95% CI) 1.99 (1.02, 3.89), P = 0.04; respectively] (Table S5). On the other hand, Dendritic_cells_activated_CIBERSORT showed a negative association with OS [HR (95% CI) 0.68 (0.48, 0.97), P = 0.03] (Table S5). Next, the relationship between the expression of ANOS1 and cell types or scores was estimated. As shown in Figures S7 and **10A**, fibroblasts by MCPcounter [30] and xCell [31], cancer-associated fibroblasts (CAFs) by EPIC [32], stromal score by estimate [33], Mesenchymal stromal cells (MSC) by xCell analysis showed a positive relationship with ANOS1 expression, which is consistent with the findings of amplification of 11q13.3 (FGF3/4/19) and enrichment of the

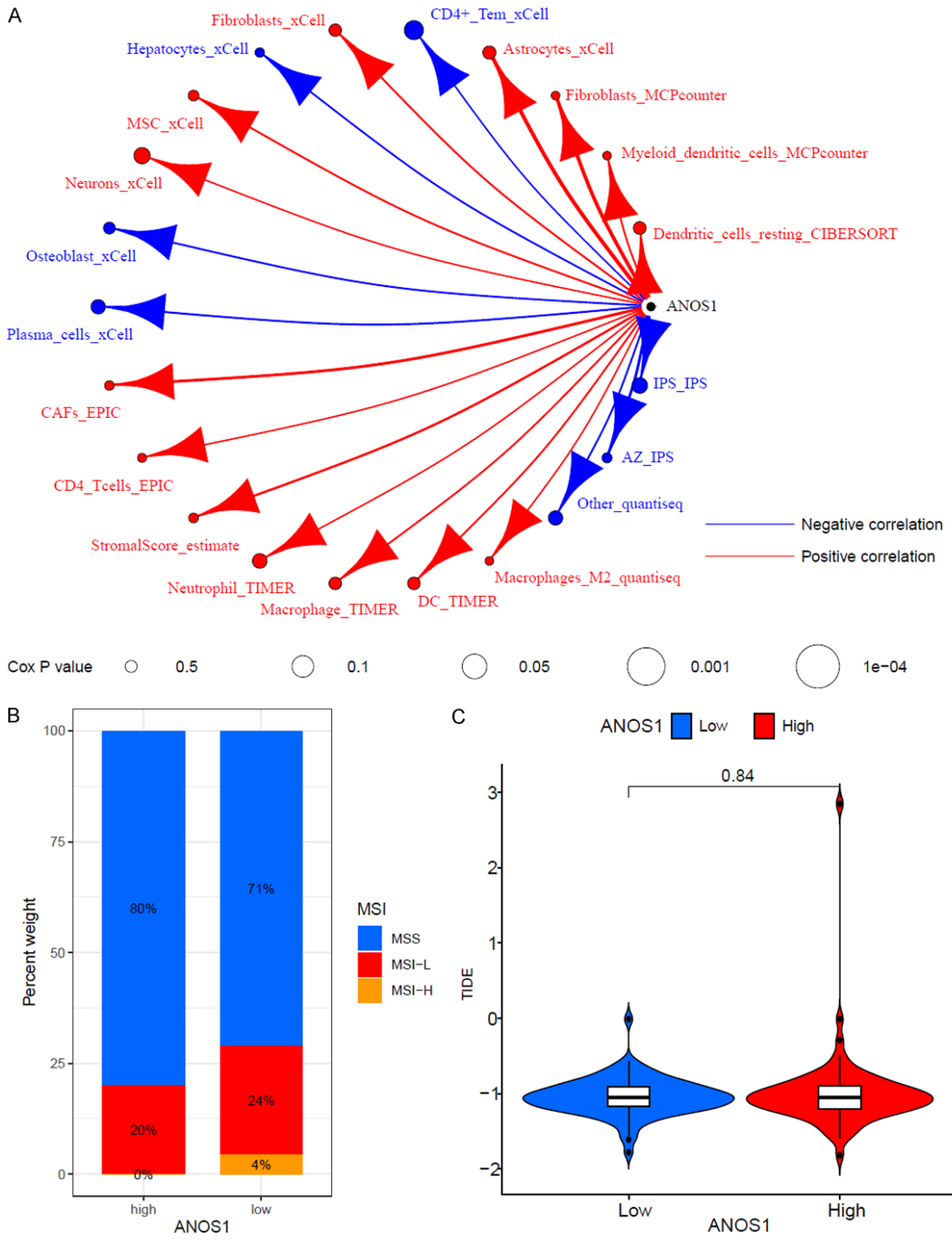


Figure 10. Tumour microenvironment analysis based on the ANOS1 expression. A. Network showing the correlation among ANOS1 expression levels and the cell types or scores calculated by different algorithms. Significantly positive and negative correlations are presented in red and blue line, respectively. The size of the nodes indicates the *p* values from Cox regression. B. Bar plots showing the proportion of the status of microsatellite instability in the low-ANOS1 and the high-ANOS1 group. C. The violin plots displaying the TIDE score concerning the ANOS1 groups.

EMT signature in the high-ANOS1 group. Moreover, ANOS1 exhibited a positive association with macrophage_TIMER and macrophages_M2, while it displayed a negative associa-

tion with IPS_IPS (immunophenoscore) [34] (**Figure 10A**) suggests the presence of an immunosuppressive TME in the high-ANOS1 group. Moreover, Dendritic_cells_resting_CIBERSORT was found to be enhanced in the high-ANOS1 group, while neutrophils_CIBERSORT were decreased (Figure S8B). However, immune function GSEA analysis revealed an enhancement in APC co-stimulation, T cell co-stimulation, and T cell co-inhibition. Therefore, further exploration is needed to understand whether ANOS1 is associated with a suppressed immune response (Figure S8A). Since MSI can serve as a predictor of the outcome of immunotherapy, the MSI status was evaluated according to the ANOS1 groups (**Figure 10B**). In the low-ANOS1 group, 24% of patients had MSI-L (microsatellite instability-low) and 4% had MSI-H (microsatellite instability-high), whereas in the high-ANOS1 group, 20% of patients had MSI-L (**Figure 10B**). Additionally, the predicted outcome of immunotherapy did not differ between the high- and low-ANOS1 groups based on the tumor immune dysfunction and exclusion (TIDE) analysis (**Figure 10C**).

Correlation analysis between the ANOS1 expression and chemotherapy sensitivity

Chemotherapy continues to play a crucial role in the treatment of ESCA patients. We examined the correlations between ANOS1 expression and the sensitivity of chemotherapeutic agents in the TCGA ESCA dataset. BMS-754807, Lisitinib, and VX-11e exhibit significantly higher IC50 values in the high ANOS1 group compared to the low ANOS1 group (Figure S9). To identify potential sensitive drugs for patients with high ANOS1, the analysis revealed that the high-ANOS1 group exhibited a lower IC50 for cisplatin, docetaxel, cyclopamine, and vinorelbine ($P < 0.05$, Figure S10). The predicted sensitivity of other chemotherapy drugs is shown in Figure S9. These findings provide a novel insight into drug sensitivity in ESCA. Moreover, the druggable targets were also assessed in Figure S5G and S5H. Notably, PIK3CA and NFE2L2 may serve as potential therapeutic targets in the high ANOS1 group (Figure S5H).

Knockdown of ANOS1 attenuates ESCA cell proliferation

We conducted an RT-qPCR assay to confirm the expression level of ANOS1 in HEEC, KYSE150, KYSE450, TE10, TE13, and Eca109 cell lines. The results demonstrated that ANOS1 was highly expressed in KYSE150 and KYSE450 cell lines compared to HEEC cells (**Figure 11A**). Therefore, KYSE150 and KYSE450 cell lines were employed for further *in vitro* experiments. Subsequently, three shRNA sequences (designated shANOS1-1, shANOS1-2, and shANOS1-3) were designed, packed into lentivirus, and used to infect KYSE150 and KYSE450 to ensure efficient knockdown. shANOS1-1 and shANOS1-2 were chosen due to their better knockdown efficiency (**Figure 11B**). As observed in **Figure 11B**, the expression of ANOS1 was significantly downregulated in KYSE150 and KYSE450 cells after sh-ANOS1 infection. CCK-8 assays showed that KYSE150 and KYSE-450 cells with sh-ANOS1 exhibited a markedly reduced growth rate compared to the negative control (**Figure 11C**). Moreover, the knockdown of ANOS1 significantly suppressed the proliferation of KYSE150 and KYSE450 cells in colony formation assays (**Figure 11D** and **11E**). KYSE150 cells were then selected for further *in vivo* experiments. As depicted in **Figure 11F** and **11G**, silencing ANOS1 led to a significant decrease in tumor growth. The tumor weight was significantly lower in the sh-ANOS1 group compared to the negative controls (**Figure 11G**). In addition, we examined the expression of anosmin-1 using IHC staining and observed its presence in both the cytoplasm and nucleus (**Figure 11H**). Subsequently, we collected 48 pairs of esophageal cancer tissues and corresponding surrounding normal esophageal tissues to compare the expression levels of anosmin-1 in the cell nucleus (**Figure 11I**). The IHC score of anosmin-1 in the nucleus was higher in the esophageal tissues than in the surrounding tissues ($P < 0.001$, **Figure 11I**). However, the IHC score for anosmin-1 in the cell nucleus exhibits limited discriminatory power between healthy esophageal tissues and esophageal cancer (AUC = 0.68, **Figure 11J**). Furthermore, we analyzed the expression of anosmin-1 in the cytoplasm using 112 cases of ESCC and 68 cases of healthy surrounding esophageal tissues, revealing a significant increase in anosmin-1 expression in the cytoplasm (**Figure 11H** and **11K**). The IHC score for anosmin-1 in the cytoplasm effectively distinguished between esophageal cancer and healthy esophageal tissues, with an AUC of 0.96 (**Figure 11L**).

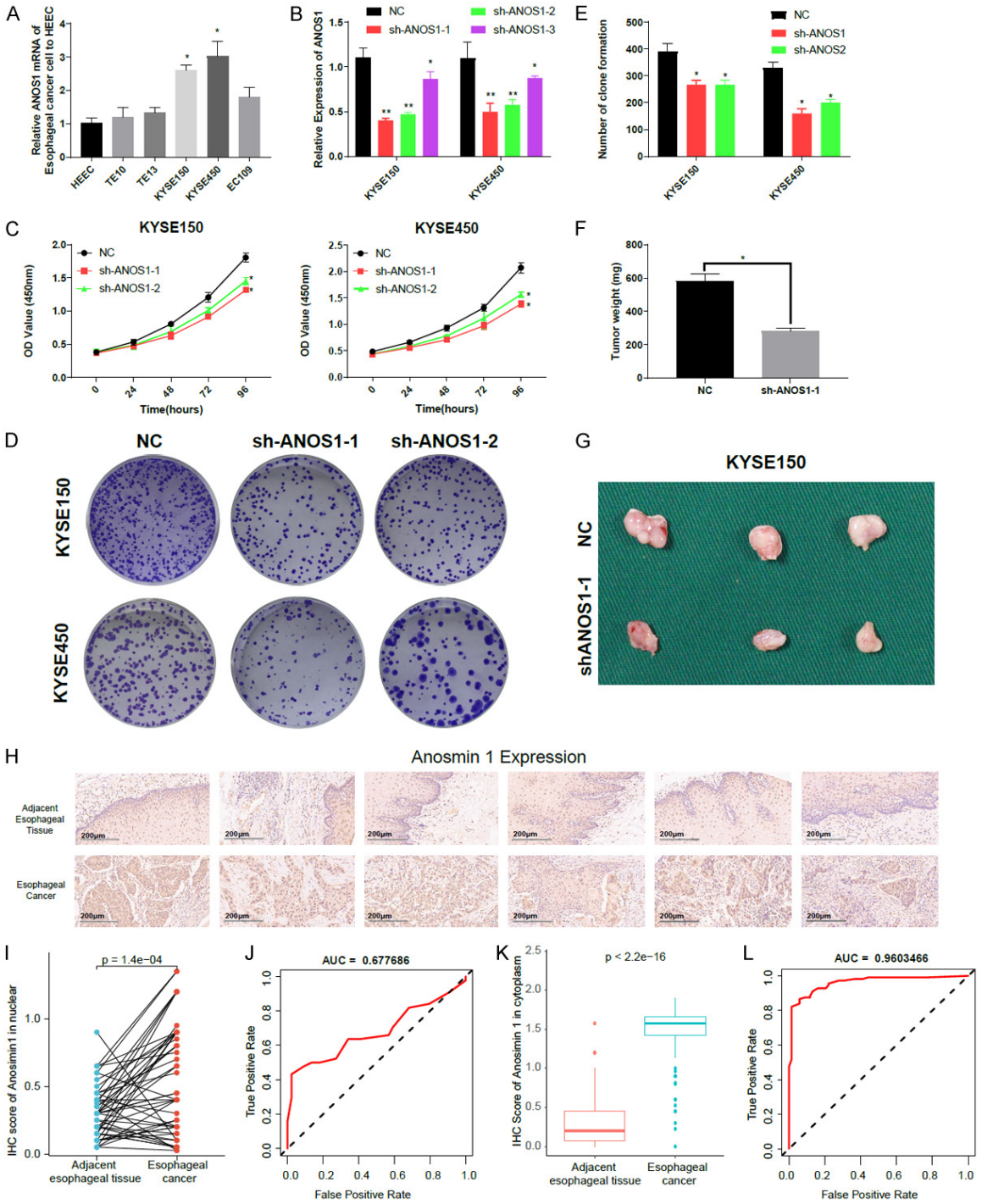


Figure 11. Results of silencing the expression of ANOS1 on ESCA cell proliferation *in vitro* and i*n vivo*. A. Differential expression of ANOS1 among ESCA cell lines and human esophageal epithelial cell line (HEECs); B. The knockdown efficiency of ANOS1 in KYSE 150 and KYSE450 using shRNA detected by RT-qPCR; C-E. The proliferation of KYSE 150 and KYSE450 cells transfected with sh-ANOS1 or relative controls measured by CCK8 assays and colony formation assays; F, G. Knockdown ANOS1 inhibits tumor growth in ESCA cancer xenograft nude mice. H. Representative immunohistochemistry (IHC) images of the expression of anosmin-1 in esophageal squamous cell carcinoma (ESCC) and adjacent esophageal tissues. I. A comparison of IHC scores for anosmin-1 in the nuclear between the ESCC and adjacent esophageal tissues. J. The ROC curve for the IHC scores of anosmin-1 in the nuclear in distinguishing between ESCC and adjacent esophageal tissues. K. A comparison of IHC scores for anosmin-1 in the cytoplasm between the ESCC and adjacent esophageal tissues. L. The ROC curve for the IHC scores of anosmin-1 in the cytoplasm in distinguishing between ESCC and adjacent esophageal tissues (*P < 0.05; **P < 0.01).

Discussion

In 2020, there were a total of 604,000 new cases of esophageal cancer, making it the seventh most common cancer worldwide [1]. Squamous cell carcinoma of the esophagus has the highest regional incidence in East Asia, particularly in China [1]. Risk factors for squamous cell carcinoma in low-income countries include betel quid chewing, consumption of pickled vegetables, consumption of very hot food and beverages, as well as heavy drinking and smoking in high-income countries [1]. On the other hand, adenocarcinoma accounts for two-thirds of all esophageal cancers in high-income countries. Its main risk factors include excess body weight, gastroesophageal reflux disease, and Barrett's esophagus [1]. Unfortunately, esophageal cancer is associated with a high mortality rate, with a total of 544,000 deaths reported in 2020, making it the sixth leading cause of cancer-related deaths [1]. Despite improvements in treatment, the 5-year overall survival rate remains around 10% [35]. The lack of early clinical symptoms often leads to a late-stage diagnosis of esophageal cancer [35]. Therefore, there is an urgent need to develop diagnostic biomarkers for early detection of esophageal carcinogenesis.

The lack of mutation of ANOS1 and its product anosmin-1 were initially identified in X-linked Kallmann syndrome (KS), which is characterized by anosmia and hypogonadotropic hypogonadism [36]. Anosmin-1 plays a crucial role in cell adhesion, cell migration, axonogenesis, and neuron migration, which helps explain the symptoms observed in KS patients [36]. More recent research has revealed that ANOS1 is abnormally overexpressed and is involved in the migration and metastasis of cancer cells. It can also serve as a biomarker for diagnosis and prognosis prediction in various types of cancer [5, 37, 38]. However, it's worth noting that ANOS1 has also been reported as a suppressor for hepatocellular cancer [7]. However, the mechanism and function of ANOS1 have not been fully characterized yet. Therefore, in our study, we utilized multi-omic data, including transcriptome, SNP, and CNV, to investigate the role of ANOS1 in ESCA.

We observed that ANOS1 expression was elevated in ESCA tissues compared to normal esophageal tissue. This finding was further validated *in vitro*, where ANOS1 expression was higher in ESCA cell lines compared to immortalized esophagus epithelial cells. ANOS1 expression could accurately differentiate between normal and cancerous esophageal tissue, with an AUC>0.7 in both TCGA and GSE53625 datasets. Furthermore, high ANOS1 expression was associated with advanced T stages and worse DFS, leading us to develop a clinically applicable monogram for predicting prognosis.

In our transcriptome analysis, we discovered that high ANOS1 expression was associated with enrichment in pathways related to EMT, hypoxia, angiogenesis, WNT signaling, and TGFβ signaling. Additionally, ANOS1 expression showed a positive correlation with several transcription factors, including SOX2, TCF7L1, and TFAP2C. Of these differentially expressed transcription factors, SOX2 plays a key role in reprogramming somatic cells into induced pluripotent stem cells [39], which is amplified in ESCA and enhances proliferation and anchorage-independent growth [40]. TFAP2C and its product, the transcription factor AP-2γ, are known to play roles in maintaining pluripotency and inhibiting somatic differentiation in germ cells [41, 42]. In breast cancer, the AP-2γ homodimer promotes tumor growth by upregulating the expression of EGFR [43, 44]. Additionally, through complex formation with KDM5B and MYC, the AP-2γ homodimer suppresses the expression of CDKN1A, facilitating cell cycle progression and carcinogenesis [45]. The increased expression of SOX2 and TFAP2C in the high-ANOS1 group suggests a potential increase in stemness. Furthermore, our analysis identified the involvement of lncRNAs in post-transcriptional regulations. For example, TUG1 has been reported to promote colony formation and invasion *in vitro*, as well as tumor growth and metastasis in oral squamous cell carcinoma *in vivo* [46, 47]. Additionally, LINC00641 was shown to regulate tumor cell proliferation, invasion, metastasis, and apoptosis [48]. These findings highlight the potential importance of lncRNAs in the context of ANOS1 and ESCA.

We further analyzed the SNP data, considering the negative association between ANOS1 expression and APOBEC1. The APOBEC family of cytosine deaminases is responsible for

endogenous mutagenesis [26, 49]. APOBEC signatures are characterized by C>T and C>G mutations in the tCw context (w = A or T) [49]. A slightly higher mutation burden was observed in the low-ANOS1 group (median tumor mutational burden [TMB] of 1.96/MB and median 98 variants per sample) compared to the high-ANOS1 group (median TMB of 1.63/MB and median 81.5 variants per sample) may be attributed to the negative association between ANOS1 expression and APOBEC1. Interestingly, the low-ANOS1 group exhibited two specific mutational signatures, SBS6 and ID2, which are associated with defective mismatch repair and microsatellite instability [25]. This finding aligns with the observation that 24% of patients in the low-ANOS1 group had MSI-L (microsatellite instability-low) and 4% had MSI-H (microsatellite instability-high).

In addition to endogenous mutagens, oxidative stress caused by gastroesophageal reflux is a well-known external mutagen that can lead to DNA damage [50-52]. The SBS17b signature, observed in both the high- and low-ANOS1 groups, is active across esophageal cancers and is associated with Barrett's esophagus or tumors exposed to Capecitabine/5-FU [53]. This signature appears early during the development of ESCA [54] and is speculated to be caused by oxidative damage resulting from gastroesophageal reflux [23, 28]. Oxidative stress induced by a low pH bile acid cocktail can oxidize dGTP to 8-oxo-dGTP, leading to T>G transversions at TT sites [28, 55].

Furthermore, signatures SBS1, SBS5, and SBS40 are likely the result of accumulated mutations over a lifetime, as they show a significant association with age at diagnosis [56]. SBS1, identified in the high-ANOS1 group, is reflective of biological cell aging [57] and is caused by the deamination of 5-methylcytosine to thymine [58]. There is a female-dominated bias in the proportion of samples positive for SBS1 [58]. SBS40, found in the low-ANOS1 group, and SBS5, found in the high-ANOS1 group, are also associated with age at diagnosis and have been identified through extensive screening of cancer samples [24, 56, 59], although their underlying causes have not yet been determined.

We further estimated the DBS and ID concerning ANOS1 expression. In both the high- and low-ANOS1 groups, we identified DBS2 and DBS4. DBS2 is associated with exposure to tobacco smoking and mutagens like acetaldehyde [24]. It exhibits transcriptional strand bias, with more GG>TT mutations than CC>AA on untranscribed gene strands, indicating damage to guanine and repair through transcription-coupled nucleotide excision repair [24]. DBS4 correlates with age at cancer diagnosis [24]. Additionally, in the high-ANOS1 group, we identified DBS1, which also shows transcriptional strand bias, with more CC>TT mutations than GG>AA on untranscribed strands, indicating damage to cytosine and repair via transcription-coupled nucleotide excision repair [24].

Furthermore, we extracted and matched the ID signatures. In the high-ANOS1 group, we found ID8 and ID9. Notably, ID9, which is associated with chromothripsis [27], has been linked to cancer samples characterized by extensive genomic rearrangements occurring in a single catastrophic event, leading to multiple genomic aberrations [60]. Chromothripsis is associated with increased cancer aggressiveness and poor prognosis [27], consistent with our findings that high ANOS1 expression is associated with advanced T stages and worse DFS. ID9 was also significantly associated with hypoxia score [61], aligning with our GSEA results showing enrichment of hypoxia in the high-ANOS1 group. The relationship between ANOS1 and chromothripsis needs further investigation. As for ID8, present in the high-ANOS1 group, it may be associated with DNA double-strand break repair through non-homologous end-joining mechanisms, leading to deletions longer than 5 base pairs in most tumors. In a small number of cases, ID8, along with ID17, may be linked to the somatic mutation p.K743N of TOP2A, resulting in deletions of 6 to 8 base pairs [24]. Regarding ID2, identified in the low-ANOS1 group, it is associated with age at diagnosis in non-hypermutated samples in a clock-like behavior [24], although it is not directly related to ANOS1.

In addition to the SNP analysis, we also explored the CNV data. As a result, we identified amplification of the 11q13.3 region in the high-ANOS1 group, which is associated with chromosomal instability [62]. This region contains oncogenes such as FGF3/4/19, CCND1, and FADD [29], which have been linked to node

metastasis in ESCC [63]. Interestingly, we observed that ANOS1 expression tended to be higher in the N2 stage compared to the N1 stage in ESCC.

Although there was no difference in the predicted outcome of immunotherapy, we found that cisplatin, docetaxel, cyclopamine, and vinorelbine exhibited higher sensitivity in the high-ANOS1 group. The high expression of ANOS1 promotes esophageal cancer progression, yet in our predictions, it appears to be sensitivity to specific chemotherapeutic drugs. Possible reasons for this contradiction may due to the interaction between the mechanisms of specific chemotherapeutic agents and the functionality of the ANOS1 gene. This seemingly paradoxical scenario is also evident in individuals carrying germ-line BRCA1 and BRCA2 mutations, with a significantly elevated lifetime risk of developing breast cancer. However, compared to breast cancer cells without BRCA1 pathogenic variants, breast cancer cells with BRCA1 pathogenic variants exhibit increased sensitivity to platinum-based chemotherapeutic agents, which disrupt DNA structure [64]. Additionally, these chemotherapeutic agents might have the capability to interfere with the expression of ANOS1, consequently leading to cellular apoptosis. Moreover, the elevated ANOS1 expression promote tumor cell growth and render them more sensitive to certain chemotherapeutic agents. More experiments are needed to validate the impact of ANOS1 expression on chemosensitivity. These findings may provide novel insights for the treatment of ESCA.

Conclusion

In summary, our study reveals that ANOS1 expression is increased in ESCA. Importantly, we demonstrate that ANOS1 accelerates the progression of ESCA through multi-omic analysis and experimental validation. While further investigation is needed to understand the pathological role of ANOS1 and its regulation during the development of ESCA, our data provide a rationale for exploring ANOS1 as a biomarker and a potential therapeutic target in ESCA.

Acknowledgements

Yanjun Zhou is supported by Wuxi Translational Medicine Research Institute (Project number LCYJ202312) and Wuxi Municipal Health Commission General Project (M202327); Ke Gu is supported by the Wuxi Taihu Lake Talent Plan, Supports for Leading Talents in Medical and Health Profession, Project Plan of Wuxi Institute of Translational Medicine (LCYJ202210), Scientific Research Project of Wuxi Commission of Health (M202041), Maternal and Child Health Research Project of Jiangsu Commission of Health (F202009), and Scientific Research Project of Jiangsu Maternal and Child Health Association (FYX202016).

Disclosure of conflict of interest

None.

Address correspondence to: Ke Gu and Yanjun Zhou, Department of Radiotherapy and Oncology, The Affiliated Hospital of Jiangnan University, Wuxi 214000, Jiangsu, China. E-mail: yourprofGUKE@126.com (KG); zyjmed@yeah.net (YJZ)

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Kim E, Koroukian S and Thomas CR Jr. Conditional survival of esophageal cancer: an analysis from the SEER registry (1988-2011). J Thorac Oncol 2015; 10: 1490-1497.

[3] Choy C and Kim SH. Biological actions and interactions of anosmin-1. Front Horm Res 2010; 39: 78-93.

[4] Gonzalez-Martinez D, Kim SH, Hu Y, Guimond S, Schofield J, Winyard P, Vannelli GB, Turnbull J and Bouloux PM. Anosmin-1 modulates fibroblast growth factor receptor 1 signaling in human gonadotropin-releasing hormone olfactory neuroblasts through a heparan sulfate-dependent mechanism. J Neurosci 2004; 24: 10384-10392.

[5] Choy CT, Kim H, Lee JY, Williams DM, Palethorpe D, Fellows G, Wright AJ, Laing K, Bridges LR, Howe FA and Kim SH. Anosmin-1 contributes to brain tumor malignancy through integrin signal pathways. Endocr Relat Cancer 2013; 21: 85-99.

[6] Jian B, Nagineni CN, Meleth S, Grizzle W, Bland K, Chaudry I and Raju R. Anosmin-1 involved in neuronal cell migration is hypoxia inducible and cancer regulated. Cell Cycle 2009; 8: 3770-3776.

[7] Tanaka Y, Kanda M, Sugimoto H, Shimizu D, Sueoka S, Takami H, Ezaka K, Hashimoto R,

Okamura Y, Iwata N, Tanaka C, Yamada S, Fujii T, Nakayama G, Koike M, Nomoto S, Fujiwara M and Kodera Y. Translational implication of Kallmann syndrome-1 gene expression in hepatocellular carcinoma. Int J Oncol 2015; 46: 2546-2554.

[8] Tang Z, Li C, Kang B, Gao G, Li C and Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45: W98-W102.

[9] Iasonos A, Schrag D, Raj GV and Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 2008; 26: 1364-1370.

[10] Yu G, Li F, Qin Y, Bo X, Wu Y and Wang S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. Bioinformatics 2010; 26: 976-978.

[11] Hanzelmann S, Castelo R and Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013; 14: 7.

[12] Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, Tang Y, Chen YG, Jin CN, Yu Y, Xu JT, Li YM, Cai XX, Zhou ZY, Chen XH, Pei YY, Hu L, Su JJ, Cui SD, Wang F, Xie YY, Ding SY, Luo MF, Chou CH, Chang NW, Chen KW, Cheng YH, Wan XH, Hsu WL, Lee TY, Wei FX and Huang HD. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res 2020; 48: D148-D154.

[13] Li JH, Liu S, Zhou H, Qu LH and Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014; 42: D92-97.

[14] Mayakonda A, Lin DC, Assenov Y, Plass C and Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 2018; 28: 1747-1756.

[15] Zhou Y, Zhu J, Gu M and Gu K. Prognosis and characterization of microenvironment in cervical cancer influenced by fatty acid metabolism-related genes. J Oncol 2023; 2023: 6851036.

[16] Wang S, Tao Z, Wu T and Liu XS. Sigflow: an automated and comprehensive pipeline for cancer genome mutational signature analysis. Bioinformatics 2021; 37: 1590-1592.

[17] Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor JF, Schrock AB, Hartmaier RJ, Trabucco SE, Gay L, Ali SM, Elvin JA, Singal G, Ross JS, Fabrizio D, Szabo PM, Chang H, Sasson A, Srinivasan S, Kirov S, Szustakowski J, Vitazka P, Edwards R, Bufill JA, Sharma N, Ou SI, Peled N, Spigel DR, Rizvi H, Aguilar EJ, Carter BW, Erasmus J, Halpenny DF, Plodkowski AJ, Long NM, Nishino M, Denning WL, Galan-Cobo A, Hamdi H, Hirz T, Tong P, Wang J, Rodriguez-Canales J, Villalobos PA, Parra ER, Kalhor N, Sholl LM, Sauter JL, Jungbluth AA, Mino-Kenudson M, Azimi R, Elamin YY, Zhang J, Leonardi GC, Jiang F, Wong KK, Lee JJ, Papadimitrakopoulou VA, Wistuba II, Miller VA, Frampton GM, Wolchok JD, Shaw AT, Janne PA, Stephens PJ, Rudin CM, Geese WJ, Albacker LA and Heymach JV. STK11/LKB1 mutations and PD-1 inhibitor resistance in KRAS-mutant lung adenocarcinoma. Cancer Discov 2018; 8: 822-835.

[18] Zeng D, Ye Z, Shen R, Yu G, Wu J, Xiong Y, Zhou R, Qiu W, Huang N, Sun L, Li X, Bin J, Liao Y, Shi M and Liao W. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. Front Immunol 2021; 12: 687975.

[19] Geeleher P, Cox N and Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One 2014; 9: e107468.

[20] Lada AG, Dhar A, Boissy RJ, Hirano M, Rubel AA, Rogozin IB and Pavlov YI. AID/APOBEC cytosine deaminase induces genome-wide kataegis. Biol Direct 2012; 7: 47; discussion 47.

[21] Martinez-Jimenez F, Muinos F, Sentis I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, Mularoni L, Pich O, Bonet J, Kranas H, Gonzalez-Perez A and Lopez-Bigas N. A compendium of mutational cancer driver genes. Nat Rev Cancer 2020; 20: 555-572.

[22] Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D, Mitchell TJ, Rubanova Y, Anur P, Yu K, Tarabichi M, Deshwar A, Wintersinger J, Kleinheinz K, Vazquez-Garcia I, Haase K, Jerman L, Sengupta S, Macintyre G, Malikic S, Donmez N, Livitz DG, Cmero M, Demeulemeester J, Schumacher S, Fan Y, Yao X, Lee J, Schlesner M, Boutros PC, Bowtell DD, Zhu H, Getz G, Imielinski M, Beroukhim R, Sahinalp SC, Ji Y, Peifer M, Markowetz F, Mustonen V, Yuan K, Wang W, Morris QD; PCAWG Evolution & Heterogeneity Working Group, Spellman PT, Wedge DC and Van Loo P; PCAWG Consortium. The evolutionary history of 2,658 cancers. Nature 2020; 578: 122-128.

[23] Dulak AM, Stojanov P, Peng S, Lawrence MS, Fox C, Stewart C, Bandla S, Imamura Y, Schumacher SE, Shefler E, McKenna A, Carter SL, Cibulskis K, Sivachenko A, Saksena G, Voet D, Ramos AH, Auclair D, Thompson K, Sougnez C, Onofrio RC, Guiducci C, Beroukhim R, Zhou Z, Lin L, Lin J, Reddy R, Chang A, Landrenau R, Pennathur A, Ogino S, Luketich JD, Golub TR, Gabriel SB, Lander ES, Beer DG, Godfrey TE, Getz G and Bass AJ. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and muta-

tional complexity. Nat Genet 2013; 45: 478-486.
[24] Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, Islam SMA, Lopez-Bigas N, Klimczak LJ, McPherson JR, Morganella S, Sabarinathan R, Wheeler DA, Mustonen V; PCAWG Mutational Signatures Working Group, Getz G, Rozen SG and Stratton MR; PCAWG Consortium. The repertoire of mutational signatures in human cancer. Nature 2020; 578: 94-101.
[25] Petljak M, Alexandrov LB, Brammeld JS, Price S, Wedge DC, Grossmann S, Dawson KJ, Ju YS, Iorio F, Tubio JMC, Koh CC, Georgakopoulos-Soares I, Rodriguez-Martin B, Otlu B, O'Meara S, Butler AP, Menzies A, Bhosle SG, Raine K, Jones DR, Teague JW, Beal K, Latimer C, O'Neill L, Zamora J, Anderson E, Patel N, Maddison M, Ng BL, Graham J, Garnett MJ, McDermott U, Nik-Zainal S, Campbell PJ and Stratton MR. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. Cell 2019; 176: 1282-1294, e1220.
[26] Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, Hinoue T, Laird PW, Hoadley KA, Akbani R, Castro MAA, Gibb EA, Kanchi RS, Gordenin DA, Shukla SA, Sanchez-Vega F, Hansel DE, Czerniak BA, Reuter VE, Su X, de Sa Carvalho B, Chagas VS, Mungall KL, Sadeghi S, Pedamallu CS, Lu Y, Klimczak LJ, Zhang J, Choo C, Ojesina AI, Bullman S, Leraas KM, Lichtenberg TM, Wu CJ, Schultz N, Getz G, Meyerson M, Mills GB and McConkey DJ; TCGA Research Network, Weinstein JN, Kwiatkowski DJ and Lerner SP. Comprehensive molecular characterization of muscle-invasive bladder cancer. Cell 2017; 171: 540-556, e525.
[27] Voronina N, Wong JKL, Hubschmann D, Hlevnjak M, Uhrig S, Heilig CE, Horak P, Kreutzfeldt S, Mock A, Stenzinger A, Hutter B, Frohlich M, Brors B, Jahn A, Klink B, Gieldon L, Sieverling L, Feuerbach L, Chudasama P, Beck K, Kroiss M, Heining C, Mohrmann L, Fischer A, Schrock E, Glimm H, Zapatka M, Lichter P, Frohling S and Ernst A. The landscape of chromothripsis across adult cancer types. Nat Commun 2020; 11: 2320.
[28] Nones K, Waddell N, Wayte N, Patch AM, Bailey P, Newell F, Holmes O, Fink JL, Quinn MCJ, Tang YH, Lampe G, Quek K, Loffler KA, Manning S, Idrisoglu S, Miller D, Xu Q, Waddell N, Wilson PJ, Bruxner TJC, Christ AN, Harliwong I, Nourse C, Nourbakhsh E, Anderson M, Kazakoff S, Leonard C, Wood S, Simpson PT, Reid LE, Krause L, Hussey DJ, Watson DI, Lord RV, Nancarrow D, Phillips WA, Gotley D, Smithers BM, Whiteman DC, Hayward NK, Campbell PJ, Pearson JV, Grimmond SM and Barbour AP. Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. Nat Commun 2014; 5: 5224.
[29] Gibcus JH, Menkema L, Mastik MF, Hermsen MA, de Bock GH, van Velthuysen ML, Takes RP, Kok K, Alvarez Marcos CA, van der Laan BF, van den Brekel MW, Langendijk JA, Kluin PM, van der Wal JE and Schuuring E. Amplicon mapping and expression profiling identify the Fas-associated death domain gene as a new driver in the 11q13.3 amplicon in laryngeal/pharyngeal cancer. Clin Cancer Res 2007; 13: 6257-6266.
[30] Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH and de Reyniès A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016; 17: 218.
[31] Aran D, Hu Z and Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017; 18: 220.
[32] Racle J, de Jonge K, Baumgaertner P, Speiser DE and Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. Elife 2017; 6: e26476.
[33] Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB and Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013; 4: 2612.
[34] Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H and Trajanoski Z. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep 2017; 18: 248-262.
[35] Huang FL and Yu SJ. Esophageal cancer: risk factors, genetic association, and treatment. Asian J Surg 2018; 41: 210-215.
[36] de Castro F, Seal R and Maggi R; Group of HGNC consultants for KAL1 nomenclature. ANOS1: a unified nomenclature for Kallmann syndrome 1 gene (KAL1) and anosmin-1. Brief Funct Genomics 2017; 16: 205-210.
[37] Tengara S, Tominaga M, Kamo A, Taneda K, Negi O, Ogawa H and Takamori K. Keratinocyte-derived anosmin-1, an extracellular glycoprotein encoded by the X-linked Kallmann syndrome gene, is involved in modulation of epidermal nerve density in atopic dermatitis. J Dermatol Sci 2010; 58: 64-71.

[38] Arikawa T, Kurokawa T, Ohwa Y, Ito N, Kotake K, Nagata H, Miyachi M, Suzumura K and Nonami T. Risk factors for surgical site infection after hepatectomy for hepatocellular carcinoma. Hepatogastroenterology 2011; 58: 143-146.
[39] Novak D, Huser L, Elton JJ, Umansky V, Altevogt P and Utikal J. SOX2 in development and cancer biology. Semin Cancer Biol 2020; 67: 74-82.
[40] Bass AJ, Watanabe H, Mermel CH, Yu S, Perner S, Verhaak RG, Kim SY, Wardwell L, Tamayo P, Gat-Viks I, Ramos AH, Woo MS, Weir BA, Getz G, Beroukhim R, O'Kelly M, Dutt A, Rozenblatt-Rosen O, Dziunycz P, Komisarof J, Chirieac LR, Lafargue CJ, Scheble V, Wilbertz T, Ma C, Rao S, Nakagawa H, Stairs DB, Lin L, Giordano TJ, Wagner P, Minna JD, Gazdar AF, Zhu CQ, Brose MS, Cecconello I, Ribeiro U Jr, Marie SK, Dahl O, Shivdasani RA, Tsao MS, Rubin MA, Wong KK, Regev A, Hahn WC, Beer DG, Rustgi AK and Meyerson M. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet 2009; 41: 1238-1242.
[41] Schafer S, Anschlag J, Nettersheim D, Haas N, Pawig L and Schorle H. The role of BLIMP1 and its putative downstream target TFAP2C in germ cell development and germ cell tumours. Int J Androl 2011; 34: e152-158; discussion e158-159.
[42] Pastor WA, Liu W, Chen D, Ho J, Kim R, Hunt TJ, Lukianchikov A, Liu X, Polo JM, Jacobsen SE and Clark AT. TFAP2C regulates transcription in human naive pluripotency by opening enhancers. Nat Cell Biol 2018; 20: 553-564.
[43] Park JM, Wu T, Cyr AR, Woodfield GW, De Andrade JP, Spanheimer PM, Li T, Sugg SL, Lal G, Domann FE, Zhang W and Weigel RJ. The role of Tcfap2c in tumorigenesis and cancer growth in an activated Neu model of mammary carcinogenesis. Oncogene 2015; 34: 6105-6114.
[44] De Andrade JP, Park JM, Gu VW, Woodfield GW, Kulak MV, Lorenzen AW, Wu VT, Van Dorin SE, Spanheimer PM and Weigel RJ. EGFR is regulated by TFAP2C in luminal breast cancer and is a target for vandetanib. Mol Cancer Ther 2016; 15: 503-511.
[45] Wong PP, Miranda F, Chan KV, Berlato C, Hurst HC and Scibetta AG. Histone demethylase KDM5B collaborates with TFAP2C and Myc to repress the cell cycle inhibitor p21(cip) (CDKN1A). Mol Cell Biol 2012; 32: 1633-1644.
[46] Yan G, Wang X, Yang M, Lu L and Zhou Q. Long non-coding RNA TUG1 promotes progression of oral squamous cell carcinoma through upregulating FMNL2 by sponging miR-219. Am J Cancer Res 2017; 7: 1899-1912.
[47] Liang S, Zhang S, Wang P, Yang C, Shang C, Yang J and Wang J. LncRNA, TUG1 regulates the oral squamous cell carcinoma progression possibly via interacting with Wnt/beta-catenin signaling. Gene 2017; 608: 49-57.
[48] Han X and Zhang S. Role of long non-coding RNA LINC00641 in cancer. Front Oncol 2022; 11: 829137.
[49] Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, Hess JM, Rheinbay E, Kim J, Maruvka YE, Braunstein LZ, Kamburov A, Hanawalt PC, Wheeler DA, Koren A, Lawrence MS and Getz G. Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. Cell 2016; 164: 538-549.
[50] Kauppi J, Rasanen J, Sihvo E, Nieminen U, Arkkila P, Ahotupa M and Salo J. Increased oxidative stress in the proximal stomach of patients with Barrett's esophagus and adenocarcinoma of the esophagus and esophagogastric junction. Transl Oncol 2016; 9: 336-339.
[51] Jimenez P, Piazuelo E, Sanchez MT, Ortego J, Soteras F and Lanas A. Free radicals and antioxidant systems in reflux esophagitis and Barrett's esophagus. World J Gastroenterol 2005; 11: 2697-2703.
[52] Dvorak K, Payne CM, Chavarria M, Ramsey L, Dvorakova B, Bernstein H, Holubec H, Sampliner RE, Guy N, Condon A, Bernstein C, Green SB, Prasad A and Garewal HS. Bile acids in combination with low pH induce oxidative stress and oxidative DNA damage: relevance to the pathogenesis of Barrett's oesophagus. Gut 2007; 56: 763-771.
[53] Pich O, Muinos F, Lolkema MP, Steeghs N, Gonzalez-Perez A and Lopez-Bigas N. The mutational footprints of cancer therapies. Nat Genet 2019; 51: 1732-1740.
[54] Tomkova M, Tomek J, Kriaucionis S and Schuster-Bockler B. Mutational signature distribution varies with DNA replication timing and strand asymmetry. Genome Biol 2018; 19: 129.
[55] Satou K, Hori M, Kawai K, Kasai H, Harashima H and Kamiya H. Involvement of specialized DNA polymerases in mutagenesis by 8-hydroxy-dGTP in human cells. DNA Repair (Amst) 2009; 8: 637-642.
[56] Zhu B, Poeta ML, Costantini M, Zhang T, Shi J, Sentinelli S, Zhao W, Pompeo V, Cardelli M, Alexandrov BS, Otlu B, Hua X, Jones K, Brodie S, Dabrowska ME, Toro JR, Yeager M, Wang M, Hicks B, Alexandrov LB, Brown KM, Wedge DC, Chanock S, Fazio VM, Gallucci M and Landi MT. The genomic and epigenomic evolutionary history of papillary renal cell carcinomas. Nat Commun 2020; 11: 3096.
[57] Landau HJ, Yellapantula V, Diamond BT, Rustad EH, Maclachlan KH, Gundem G, Medina-Martinez J, Ossa JA, Levine MF, Zhou Y, Kappagantula R, Baez P, Attiyeh M, Makohon-Moore

A, Zhang L, Boyle EM, Ashby C, Blaney P, Patel M, Zhang Y, Dogan A, Chung DJ, Giralt S, Lahoud OB, Peled JU, Scordo M, Shah G, Hassoun H, Korde NS, Lesokhin AM, Lu S, Mailankody S, Shah U, Smith E, Hultcrantz ML, Ulaner GA, van Rhee F, Morgan GJ, Landgren O, Papaemmanuil E, Iacobuzio-Donahue C and Maura F. Accelerated single cell seeding in relapsed multiple myeloma. Nat Commun 2020; 11: 3617.

[58] Li CH, Prokopec SD, Sun RX, Yousif F and Schmitz N; PCAWG Tumour Subtypes and Clinical Translation, Boutros PC; PCAWG Consortium. Sex differences in oncogenic mutational processes. Nat Commun 2020; 11: 4330.

[59] Olafsson S, McIntyre RE, Coorens T, Butler T, Jung H, Robinson PS, Lee-Six H, Sanders MA, Arestang K, Dawson C, Tripathi M, Strongili K, Hooks Y, Stratton MR, Parkes M, Martincorena I, Raine T, Campbell PJ and Anderson CA. Somatic evolution in non-neoplastic IBD-affected colon. Cell 2020; 182: 672-684, e611.

[60] ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Nature 2020; 578: 82-93.

[61] Bhandari V, Li CH, Bristow RG and Boutros PC; PCAWG Consortium. Divergent mutational processes distinguish hypoxic and normoxic tumours. Nat Commun 2020; 11: 737.

[62] Watkins TBK, Lim EL, Petkovic M, Elizalde S, Birkbak NJ, Wilson GA, Moore DA, Gronroos E, Rowan A, Dewhurst SM, Demeulemeester J, Dentro SC, Horswell S, Au L, Haase K, Escudero M, Rosenthal R, Bakir MA, Xu H, Litchfield K, Lu WT, Mourikis TP, Dietzen M, Spain L, Cresswell GD, Biswas D, Lamy P, Nordentoft I, Harbst K, Castro-Giner F, Yates LR, Caramia F, Jaulin F, Vicier C, Tomlinson IPM, Brastianos PK, Cho RJ, Bastian BC, Dyrskjot L, Jonsson GB, Savas P, Loi S, Campbell PJ, Andre F, Luscombe NM, Steeghs N, Tjan-Heijnen VCG, Szallasi Z, Turajlic S, Jamal-Hanjani M, Van Loo P, Bakhoum SF, Schwarz RF, McGranahan N and Swanton C. Pervasive chromosomal instability and karyotype order in tumour evolution. Nature 2020; 587: 126-132.

[63] Ying J, Shan L, Li J, Zhong L, Xue L, Zhao H, Li L, Langford C, Guo L, Qiu T, Lu N and Tao Q. Genome-wide screening for genetic alterations in esophageal cancer by aCGH identifies 11q13 amplification oncogenes associated with nodal metastasis. PLoS One 2012; 7: e39797.

[64] Lee A, Moon BI and Kim TH. BRCA1/BRCA2 pathogenic variant breast cancer: treatment and prevention strategies. Ann Lab Med 2020; 40: 114-121.

Table S1. The clinicopathologic characteristics of ESCA patients concerning *ANOS1* expression

	High-*ANOS1*	Low-*ANOS1*	*p* Value
Case number	76	77	
Age ≥ 60 (%)	32 (42.1)	44 (57.1)	0.089
Gender (%)			0.527
Female	10 (13.2)	14 (18.2)	
Male	66 (86.8)	63 (81.8)	
Race (%)			0.007
Asian	31 (40.8)	13 (16.9)	
Black	1 (1.3)	2 (2.6)	
White	38 (50.0)	48 (62.3)	
Unknown	6 (7.9)	14 (18.2)	
Histologic type (%)			< 0.001
Squamous cell carcinoma	60 (78.9)	25 (32.5)	
Adenocarcinoma	16 (21.1)	52 (67.5)	
Tumor location (%)			0.003
Lower third of esophagus	39 (51.3)	61 (79.2)	
Middle third of esophagus	28 (36.8)	13 (16.9)	
Upper third of esophagus	3 (3.9)	2 (2.6)	
Unknown	6 (7.9)	1 (1.3)	
T stage (%)			< 0.001
T1	4 (5.3)	23 (29.9)	
T2	23 (30.3)	18 (23.4)	
T3	45 (59.2)	36 (46.8)	
T4	4 (5.3)	0 (0.0)	
N stage (%)			0.189
N0	36 (47.4)	33 (42.9)	
N1	26 (34.2)	37 (48.1)	
N2	6 (7.9)	5 (6.5)	
N3	6 (7.9)	2 (2.6)	
Unknown	2 (2.6)	0 (0.0)	
M stage (%)			0.567
M0	65 (85.5)	63 (81.8)	
M1	5 (6.6)	4 (5.2)	
Unknown	6 (7.9)	10 (13.0)	

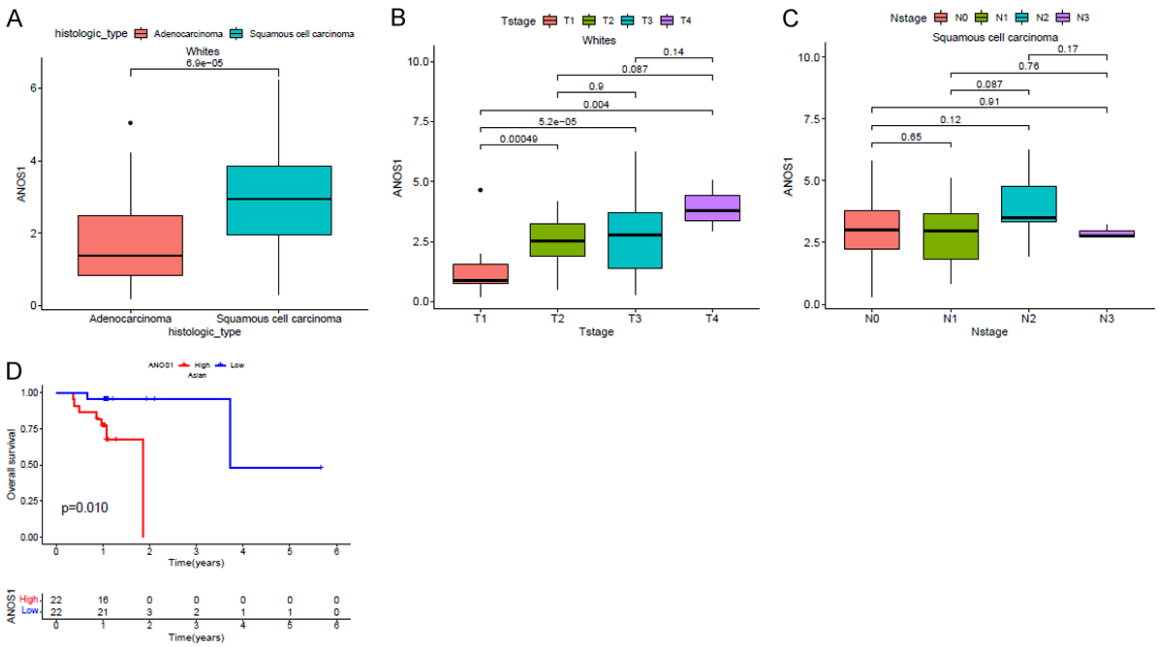


Figure S1. Association of *ANOS1* with clinicopathological parameters in the TCGA ESCA patients. Comparison of *ANOS1* expression in white ESCA patients according to histologic subtypes (A) and T stages (B). Comparison of *ANOS1* expression in squamous cell carcinoma according to N stages (C). Comparison of overall survival based on the expression of *ANOS1* in Asian patients (D).

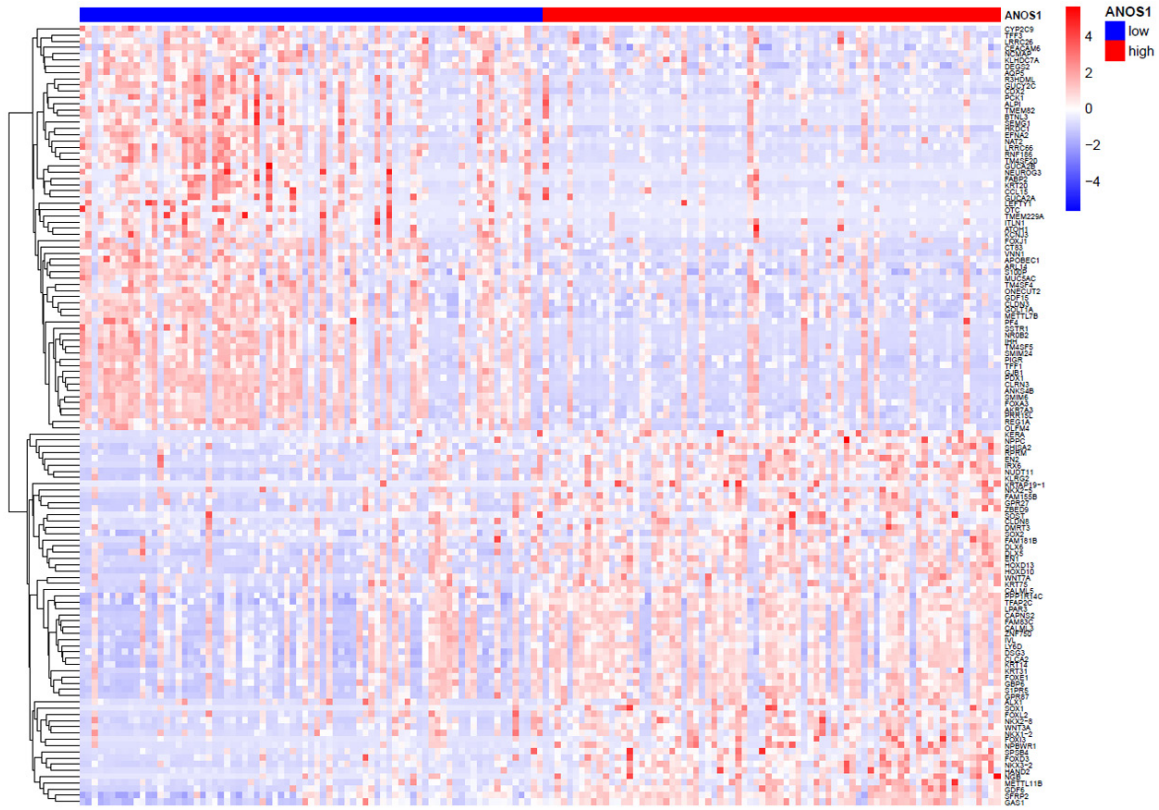


Figure S2. Heatmap of 51 differentially expressed genes concerning the *ANOS1* groups.

Table S2. The intersection of the hub genes in the "GOSemSim" and "Degree" algorithms

Only Degree	Degree AND GOSemSim
NUDT11	SFRP2
FOXI3	NKX3-2
SPSB4	GAS1
GPR27	IRX6
KERA	NPBWR1
KLRG2	EN1
RPRM	HAND2
METTL11B	TFAP2C
GDF6	LPAR3
SHISA2	EN2
NGB	FOXD3
PPP1R14C	HOXD13
FAM181B	DLX6
KRTAP19-1	WNT3A
S1PR5	FOXL2
FAM155B	DMRT3
ZBED9	SOX1
CLCA2	HOXD10
C11ORF87	NPPC
NKX2-8	DLX5
NKX1-2	FOXE1
KRT75	CALML3
ZNF750	CAPNS2
GBP6	GPR87
FAM83C	CALML5
LY6D	WNT7A
RNF186	ALX1
TMEM82	KRT31
SMIM6	KRT14
HKDC1	NKX2-5
DEGS2	SOST
CCL15	CLDN8
CLRN3	SOX2
BTNL3	DSG3
ANKS4B	IVL
GJB1	AKR7A3
GOLT1A	FOXA3
METTL7B	APOBEC1
LRRC66	IHH
SSTR1	TM4SF5
CT83	GUCA2B
TM4SF20	NR0B2
R3HDML	PDX1
SMIM24	ALPI
TMEM229A	GUCY2C
C9ORF152	FOXJ1
C2ORF72	FABP2

OTC	NAT2
PF4	KRT20
PIGR	CDX2
GDF15	CLDN3
PRR15L	LEFTY1
NCMAP	ATOH1
AC020914.1	TFF3
S100P	TFF1
SEMG1	TM4SF4
KLHDC7A	NEUROG3
ARL14	CYP2C9
VNN1	REG1A
KCNJ3	EFNA2
LRRC26	AQP5
ONECUT2	ITLN1
C11ORF86	GUCA2A
	MUC5AC
	OLFM4
	CEACAM6
	PCK1

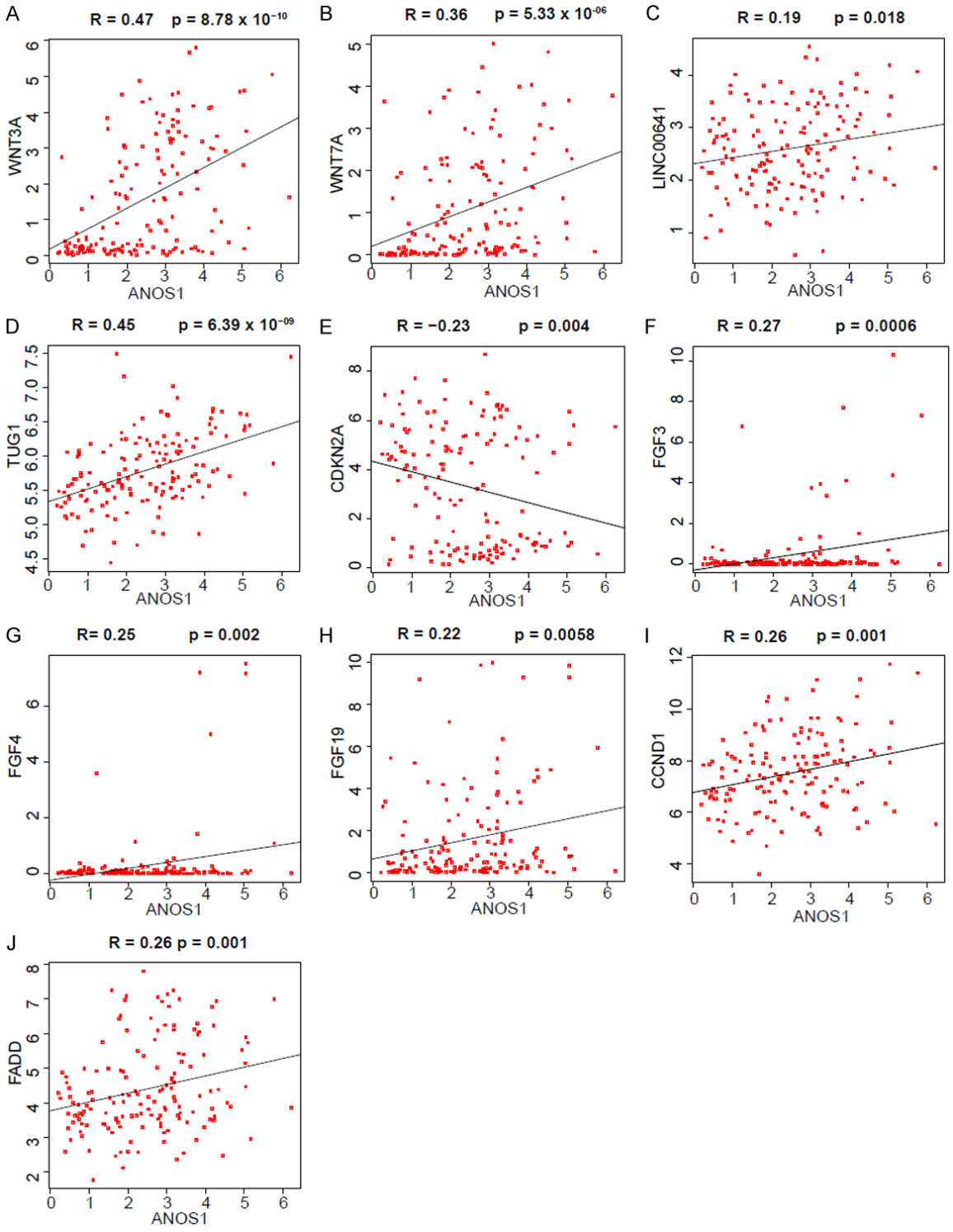
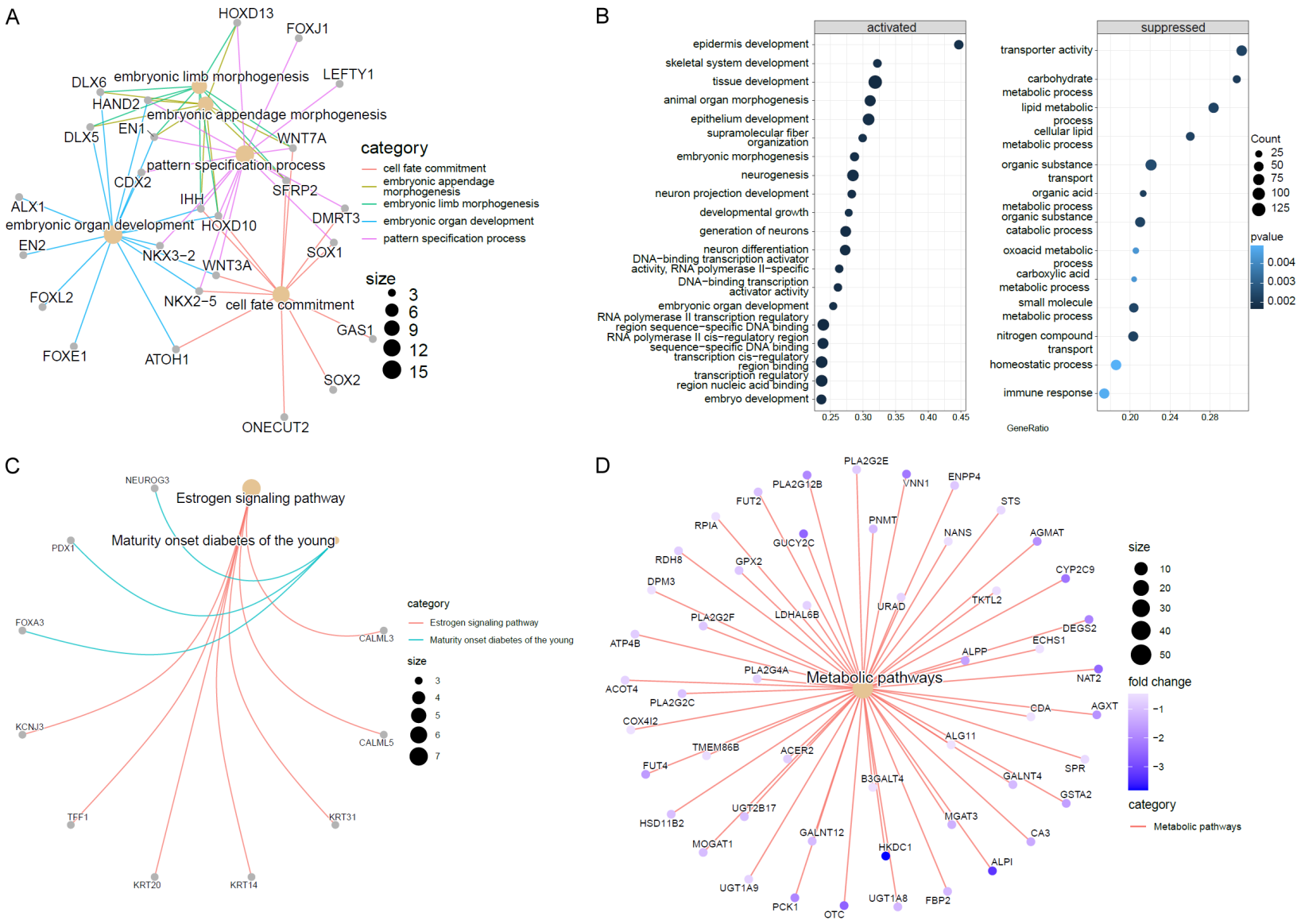


Figure S3. The correlation between the expression of *ANOS1* and other genes. The correlation between the expression of *ANOS1* and *WNT3A* (A), *WNT7A* (B), *LINC00641* (C), *TUG1* (D), *CDKN2A* (E), *FGF3* (F), *FGF4* (G), *FGF19* (H), *CCND1* (I), and *FADD* (J).

ANOS1 promotes progressions of ESCA



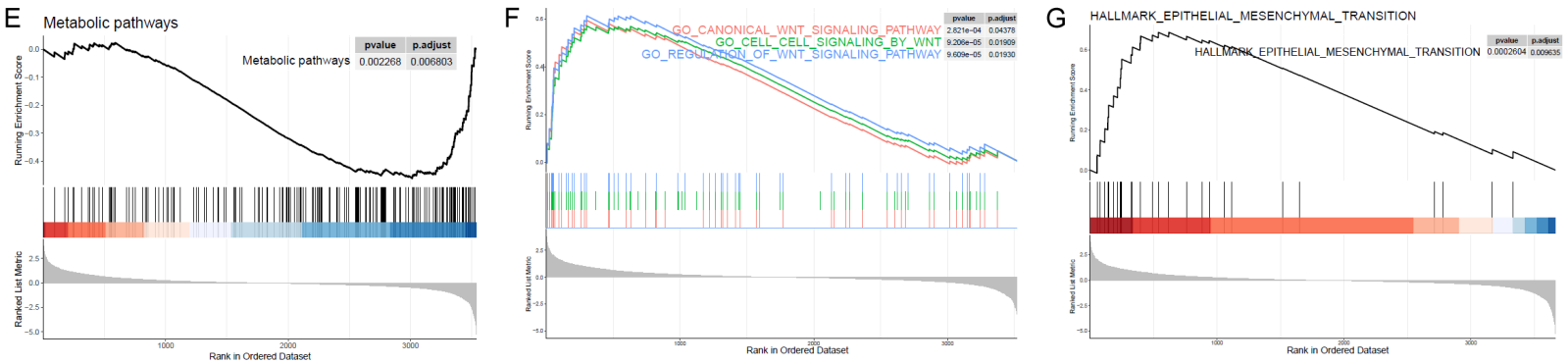


Figure S4. The functional enrichment analysis based on DEGs in ESCA. The low-ANOS1 and high- ANOS1 groups exhibits different biological function by ORA (A) and by GSEA (B), KEGG pathway by ORA (C), and by GSEA (D), metabolic pathway by GSEA (E), WNT pathway in MsigDB by GSEA (F) and EMT in HALLMARK by GSEA (G).

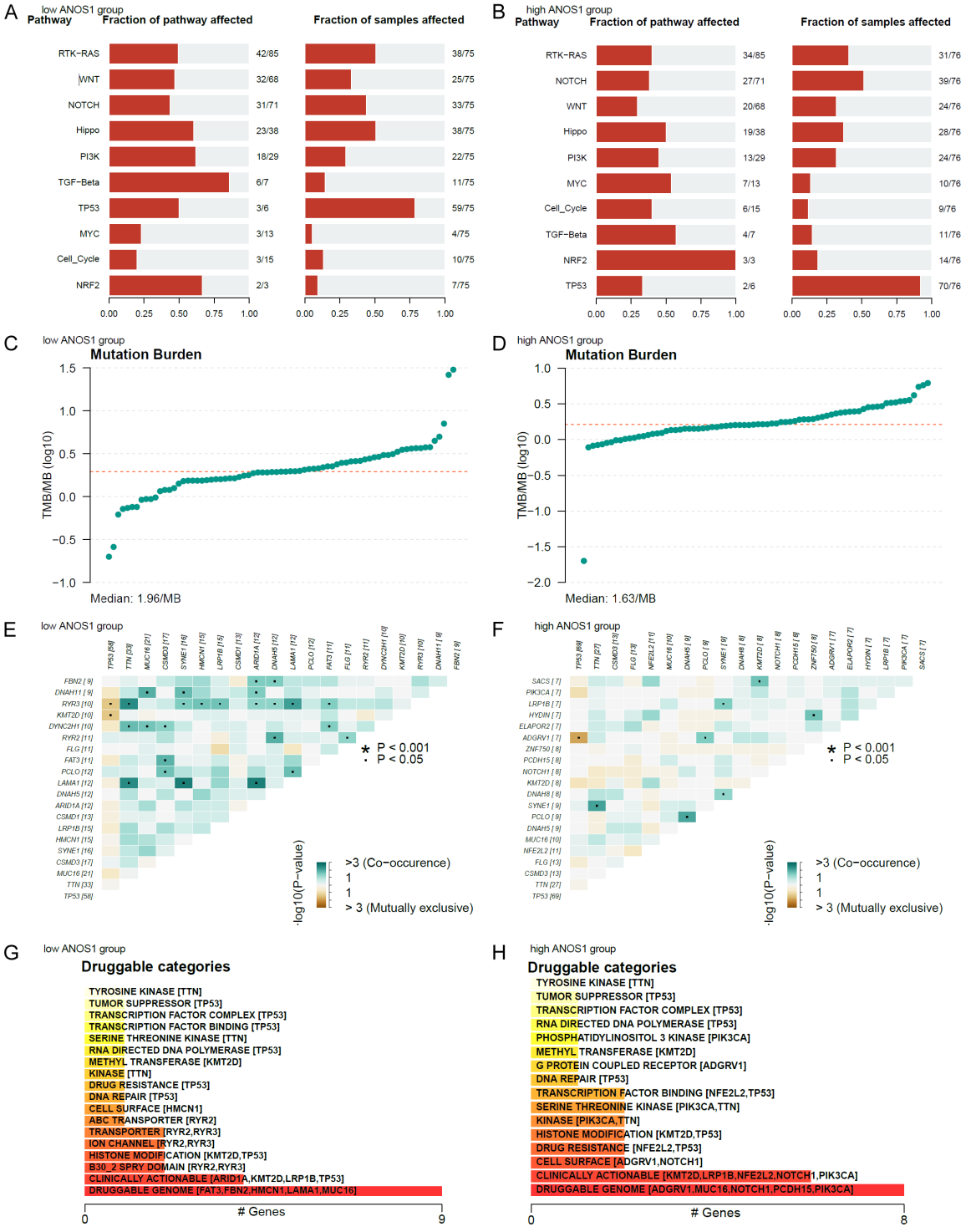


Figure S5. Single-nucleotide polymorphism analysis concerning *ANOS1* group. The enrichment of oncogenic signaling pathways in the low-*ANOS1* group (A) and in the high-*ANOS1* group (B) displayed in fraction of pathway affected and fraction of samples affected. The tumour mutation burden of each patient in the low-*ANOS1* group (C) and in the high-*ANOS1* group (D). Mutually exclusive and co-occurring gene pairs in the low-*ANOS1* group (E) and in the high-*ANOS1* group (F) displayed in a triangular matrix. Asterisk stands for P < 0.001 and dot for P < 0.05. The potential drug-target categories based on single-nucleotide polymorphism in the low-*ANOS1* group (G) and in the high-*ANOS1* group (H).

Table S3. The exacted DBS signatures matched against the COSMIC signatures

High-*ANOS1* group				Low-*ANOS1* group			
Exacted signatures	COSMIC DBS signatures	Aetiology	Similarity	Exacted signatures	COSMIC DBS signatures	Aetiology	Similarity
Sig1	DBS4	Unknown	0.642	Sig1	DBS4	Unknown	0.642
Sig2	DBS6	Unknown	0.436	Sig2	DBS4	Unknown	0.703
Sig3	DBS6	Unknown	0.167	Sig3	DBS11	Possibly related to APOBEC mutagenesis	0.424
Sig4	DBS1	Ultraviolet light exposure	0.666	Sig4	DBS8	Unknown	0.229
Sig5	DBS2	Tobacco smoking and other mutagens	0.626	Sig5	DBS2	Tobacco smoking and other mutagens	0.966

Table S4. The exacted ID signatures matched against the COSMIC signatures

High-*ANOS1* group			Low-*ANOS1* group		
Exacted signatures	COSMIC DBS signatures	Similarity	Exacted signatures	COSMIC DBS signatures	Similarity
Sig1	ID14	0.238	Sig1	ID2	0.809
Sig2	ID9	0.758	Sig2	ID8	0.39
Sig3	ID8	0.313	Sig3	ID3	0.53
Sig4	ID8	0.742	Sig4	ID11	0.467

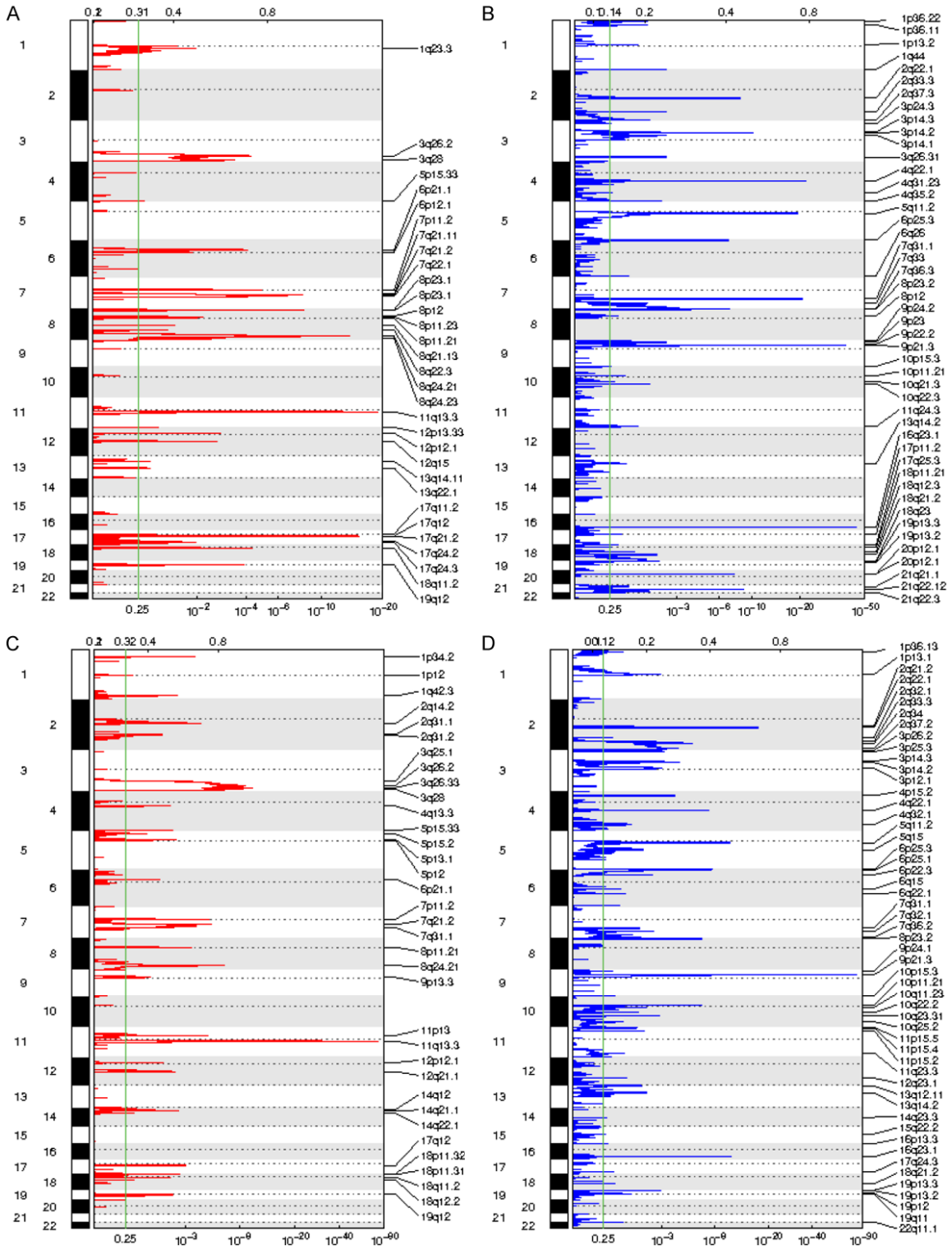


Figure S6. Genome-wide distribution of chromosome amplification and deletion. Genome-wide distribution of chromosome amplification in the low-*ANOS1* group (A), deletion in the low-*ANOS1* group (B), amplification in the high-*ANOS1* group (C), deletion in the high-*ANOS1* group (D).

Table S5. The correlation between cell types or scores and OS by univariable Cox regression

ID	Hazard_ratio	Low_CI	Ligh_CI	*p*_value
Eosinophils_CIBERSORT	1.847598797	1.159341303	2.944448977	0.009829426
Dendritic_cells_activated_CIBERSORT	0.681124915	0.480005558	0.966512037	0.031497253
Tgd_cells_xCell	1.994187034	1.02339103	3.885887027	0.042570993
CD4+_Tem_xCell	1.33442463	0.999408748	1.78174255	0.050470541
T_cells_follicular_helper_CIBERSORT	1.306668441	0.994852491	1.716216656	0.054497128
Macrophages_EPIC	1.284582116	0.990724577	1.665600362	0.058804848
Erythrocytes_xCell	1.530156503	0.980125667	2.38885584	0.061256995
Monocytic_lineage_MCPcounter	1.260103032	0.98331606	1.614800894	0.067698734
pDC_xCell	1.426653273	0.969276811	2.099853767	0.071586905
CLP_xCell	1.284455651	0.963873961	1.711661884	0.087487816
Neutrophils_xCell	1.373223503	0.950397683	1.984161813	0.091213339
Chondrocytes_xCell	0.694440409	0.452410083	1.065952107	0.095347294
CP_IPS	0.808953577	0.629683897	1.039260957	0.097179615
Neurons_xCell	1.601309552	0.89904347	2.852133812	0.109906122
CD4+_memory_T-cells_xCell	1.613724315	0.894911267	2.909904321	0.111641674
MHC_IPS	1.236171385	0.943087034	1.62033793	0.12464365
IPS_IPS	1.260065144	0.934409277	1.699217042	0.129704088
EC_IPS	1.211905719	0.940093066	1.562308588	0.138016614
MPP_xCell	2.171898663	0.776774239	6.072734609	0.139287775
Skeletal_muscle_xCell	1.274843515	0.913143065	1.779815289	0.15379159
CD8+_naive_T-cells_xCell	1.25999293	0.915188075	1.734705932	0.15657467
T_cells_MCPcounter	1.19677091	0.932290419	1.536281593	0.158620078
Neutrophil_TIMER	1.188224369	0.927763942	1.52180645	0.171919701
T_cell_CD8_TIMER	1.198028178	0.923921369	1.553456346	0.17287549
naive_B-cells_xCell	1.415499224	0.858811693	2.333035367	0.172896276
Other_quantiseq	0.808237136	0.589577995	1.107991264	0.185901216
Plasma_cells_xCell	1.240585846	0.90123233	1.707720851	0.186108799
Preadipocytes_xCell	0.797382383	0.569242613	1.116955495	0.187926643
Mast_cells_activated_CIBERSORT	1.209655566	0.900934851	1.624164707	0.205496045
ImmuneScore_estimate	1.171668918	0.91056984	1.507636199	0.218080687
Th2_cells_xCell	1.166702755	0.901630404	1.509704322	0.240997314
CD8_T_cells_MCPcounter	1.175090233	0.88916258	1.552963526	0.256721959
Astrocytes_xCell	0.826595733	0.593430148	1.151374779	0.260034643
T_cells_CD8_quantiseq	1.246647254	0.848232667	1.832197032	0.261802697
Macrophages_xCell	1.190892645	0.864430886	1.64064625	0.285184874
Dendritic_cells_resting_CIBERSORT	1.194509518	0.86040683	1.658346888	0.288334525
Macrophages_M2_xCell	1.203342543	0.849624757	1.704320955	0.297260942
DC_TIMER	1.148156715	0.885422327	1.488853175	0.29737182
Macrophage_TIMER	1.145003395	0.883344274	1.484169664	0.306346379
Macrophages_M1_xCell	1.155759808	0.874683321	1.52715926	0.308587001
Keratinocytes_xCell	0.869773528	0.663518797	1.140142516	0.312359961
Macrophages_M1_quantiseq	1.2122802	0.833444087	1.763313587	0.313954665
Fibroblasts_xCell	0.816473298	0.549689147	1.212737509	0.315158701
Megakaryocytes_xCell	0.852288796	0.622229487	1.167408822	0.319400353
aDC_xCell	1.142412246	0.877446884	1.487390021	0.322708775
T_cells_CD4_naive_CIBERSORT	1.578361016	0.613024946	4.063820752	0.344237477
Neutrophils_MCPcounter	1.144990828	0.863971636	1.51741555	0.3460217

NK_cells_MCPcounter	1.153241755	0.85312562	1.558934013	0.353884745
Th1_cells_xCell	0.860421856	0.623481398	1.187406348	0.360320064
T_cell_CD4_TIMER	1.125871064	0.873135825	1.451762276	0.360697555
ESTIMATEScore_estimate	1.120122716	0.878160226	1.428753957	0.360933312
Endothelial_cells_MCPcounter	1.128014686	0.869755987	1.46295875	0.363849679
SC_IPS	0.896303909	0.704397417	1.140493531	0.373165707
Osteoblast_xCell	0.855063422	0.604937192	1.208610524	0.375167275
Macrophages_M0_CIBERSORT	0.86030042	0.615085477	1.20327473	0.37940055
NK_cells_activated_CIBERSORT	0.861954615	0.61484539	1.208378188	0.388774067
CD8+_T-cells_xCell	1.226422952	0.768705796	1.956682601	0.391816885
Adipocytes_xCell	0.747951796	0.378103454	1.479573599	0.404053019
Macrophages_M1_CIBERSORT	1.116178112	0.856091363	1.455281097	0.416778309
otherCells_EPIC	0.899460362	0.686181319	1.179030849	0.442889441
T_cells_CD4_quantiseq	0.784114765	0.419756038	1.464745969	0.445579583
Mast_cells_resting_CIBERSORT	0.883869528	0.636354198	1.22765803	0.46148342
CD8_Tcells_EPIC	1.117557483	0.830759705	1.503364596	0.462607699
Tregs_xCell	1.11376918	0.833973776	1.487435004	0.4654016
Platelets_xCell	0.865167264	0.58272926	1.284497701	0.472581614
Monocytes_CIBERSORT	0.893877622	0.651516068	1.226396774	0.486904984
Neutrophils_CIBERSORT	1.119314206	0.814441858	1.538310291	0.48719011
CMP_xCell	0.862429569	0.564267546	1.318142019	0.494111674
Neutrophils_quantiseq	1.108966846	0.821317724	1.497358975	0.499608575
MSC_xCell	0.895408972	0.648231916	1.236837014	0.50267035
CD4+_Tcm_xCell	1.187011518	0.717687441	1.963245087	0.504256571
NK_cells_resting_CIBERSORT	1.103708108	0.817943817	1.489309612	0.518636334
CD4+_naive_T-cells_xCell	0.841996766	0.494947815	1.432390512	0.525818571
Tregs_quantiseq	1.090140683	0.832162714	1.428094156	0.531031222
Monocytes_xCell	1.107100247	0.803805648	1.524834964	0.533353229
TumorPurity_estimate	0.922893099	0.709586745	1.200320719	0.549590044
Mesangial_cells_xCell	1.083767853	0.829054539	1.416737625	0.556196826
Endothelial_EPIC	1.077025102	0.832231288	1.393822951	0.572731464
NK_cells_quantiseq	0.903589506	0.631373244	1.293171675	0.579379237
CD4+_T-cells_xCell	1.129885098	0.73213618	1.743719776	0.581220814
CD8+_Tem_xCell	1.131776851	0.728356653	1.758642331	0.58199674
NK_cells_xCell	0.814220069	0.384367022	1.724795012	0.591515191
Cytotoxic_lymphocytes_MCPcounter	1.073934466	0.817585482	1.410660123	0.608227449
ImmuneScore_xCell	1.11616552	0.729973182	1.706672929	0.611985367
AZ_IPS	1.07435251	0.811609032	1.422154349	0.616227567
iDC_xCell	0.913278701	0.640295208	1.302645992	0.616597972
B_lineage_MCPcounter	1.067612954	0.821452323	1.387539346	0.624676874
Bcells_EPIC	1.068040153	0.815705863	1.39843272	0.632171032
NKT_xCell	1.060301509	0.824833974	1.362988583	0.647677961
Sebocytes_xCell	0.938546017	0.713624762	1.234358268	0.650030882
Myocytes_xCell	0.920120015	0.640568025	1.321672032	0.652307643
DC_xCell	1.071785697	0.787583005	1.458544145	0.659215233
B_cells_quantiseq	1.091702093	0.732771199	1.62644692	0.666211206
StromalScore_estimate	1.05484854	0.827222905	1.345109565	0.666797933
Dendritic_cells_quantiseq	1.07598032	0.760595553	1.522140965	0.679039342
CAFs_EPIC	0.949586489	0.731600497	1.232523082	0.697451386

Mast_cells_xCell	0.931879941	0.651887863	1.332131297	0.698775421
Melanocytes_xCell	1.062263198	0.770884123	1.463777848	0.711949146
Hepatocytes_xCell	1.098816018	0.656653521	1.838711898	0.719785323
cDC_xCell	1.054902258	0.772782302	1.440015865	0.736407703
Smooth_muscle_xCell	0.952333962	0.714768155	1.268858958	0.738695531
CD8+_Tcm_xCell	1.065264537	0.71182263	1.59420126	0.758563646
Eosinophils_xCell	0.960123305	0.72760344	1.266949427	0.773639589
MEP_xCell	1.039595157	0.796237679	1.357331006	0.775351893
NKcells_EPIC	0.915202583	0.476599317	1.757442236	0.790103176
HSC_xCell	0.956953128	0.683017322	1.340755586	0.798161128
CD4_Tcells_EPIC	1.036733543	0.786115483	1.367250057	0.798331447
Endothelial_cells_xCell	1.041691875	0.760723914	1.426433352	0.798962261
T_cells_CD8_CIBERSORT	0.961868847	0.704357301	1.313526074	0.80680999
B_cell_TIMER	0.963556881	0.706653615	1.313857092	0.814484418
B_cells_naive_CIBERSORT	1.027982718	0.784581472	1.346894498	0.841332564
Pericytes_xCell	0.973320645	0.745544124	1.2706868	0.842416719
B_cells_memory_CIBERSORT	1.042012039	0.693512972	1.565636308	0.842956892
B-cells_xCell	1.046209896	0.663816061	1.648883192	0.845686853
MicroenvironmentScore_xCell	1.036553168	0.719463542	1.493393906	0.847193737
Basophils_xCell	1.035568656	0.724608267	1.479975442	0.847865588
Class-switched_memory_B-cells_xCell	0.956529091	0.598558775	1.528584895	0.852587784
StromaScore_xCell	0.967858262	0.677293212	1.383078405	0.857648192
Plasma_cells_CIBERSORT	0.974504888	0.725046448	1.309791639	0.864080472
T_cells_regulatory_(Tregs)_CIBERSORT	1.022737052	0.784877795	1.332680177	0.867791515
Macrophages_M2_quantiseq	0.980991249	0.718290008	1.339770596	0.903943607
Fibroblasts_MCPcounter	0.983669018	0.75037944	1.289487271	0.905106104
Myeloid_dendritic_cells_MCPcounter	1.013498704	0.784779061	1.30887746	0.918159813
Monocytes_quantiseq	1.059381352	0.324927526	3.453966681	0.923786518
T_cells_CD4_memory_resting_CIBERSORT	1.010497303	0.746760983	1.367378349	0.946048003
pro_B-cells_xCell	0.983223699	0.53643312	1.802142349	0.956354544
GMP_xCell	0.994420671	0.723581265	1.366636366	0.972486679
Memory_B-cells_xCell	0.990356658	0.570640179	1.718782424	0.972518568
Macrophages_M2_CIBERSORT	0.995559817	0.743810152	1.332516568	0.976131633
mv_Endothelial_cells_xCell	1.003715449	0.722900937	1.393613772	0.982330232
Epithelial_cells_xCell	0.999154983	0.733357989	1.361286977	0.995725462
T_cells_gamma_delta_CIBERSORT	0.000735444	0	Inf	0.996777244
T_cells_CD4_memory_activated_CIBERSORT	0.99952078	0.744277067	1.342298229	0.997457817
Endothelial_cells_xCell	1.000133135	0.697970837	1.433106132	0.999421243

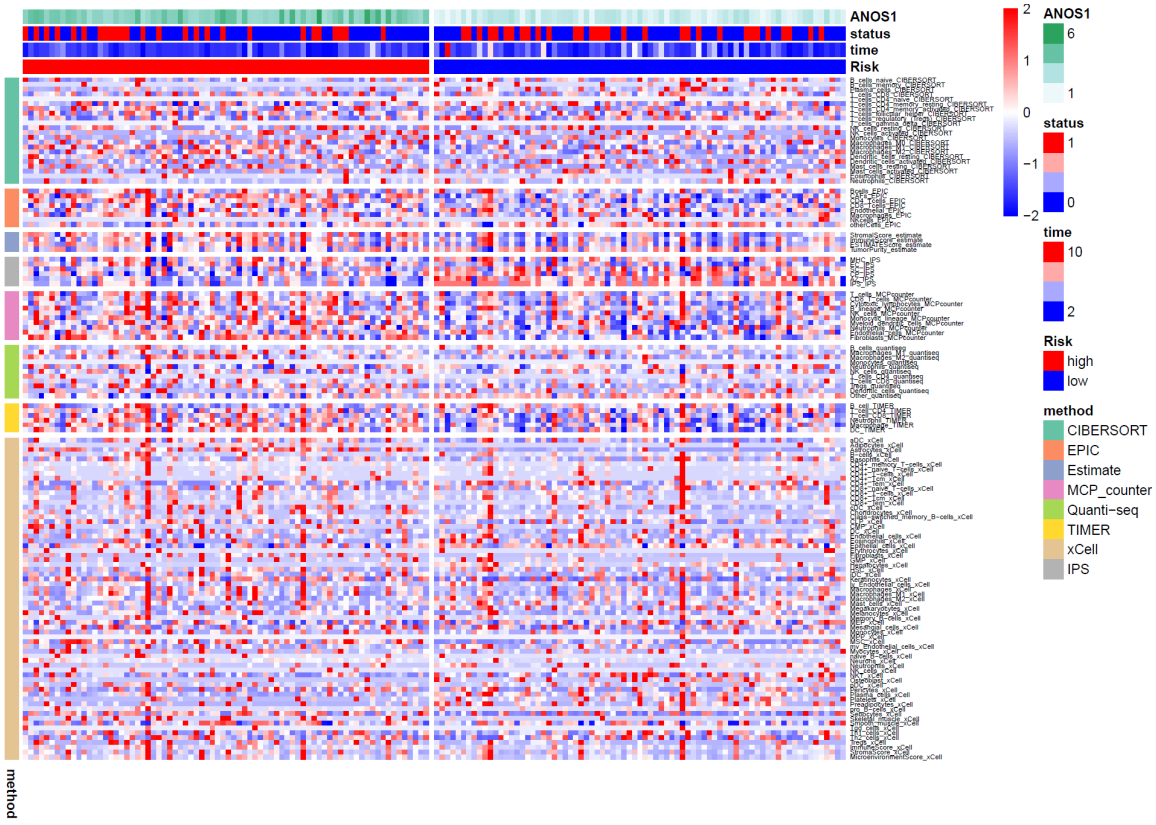


Figure S7. The correlation between the *ANOS1* and cell types by different algorithms.

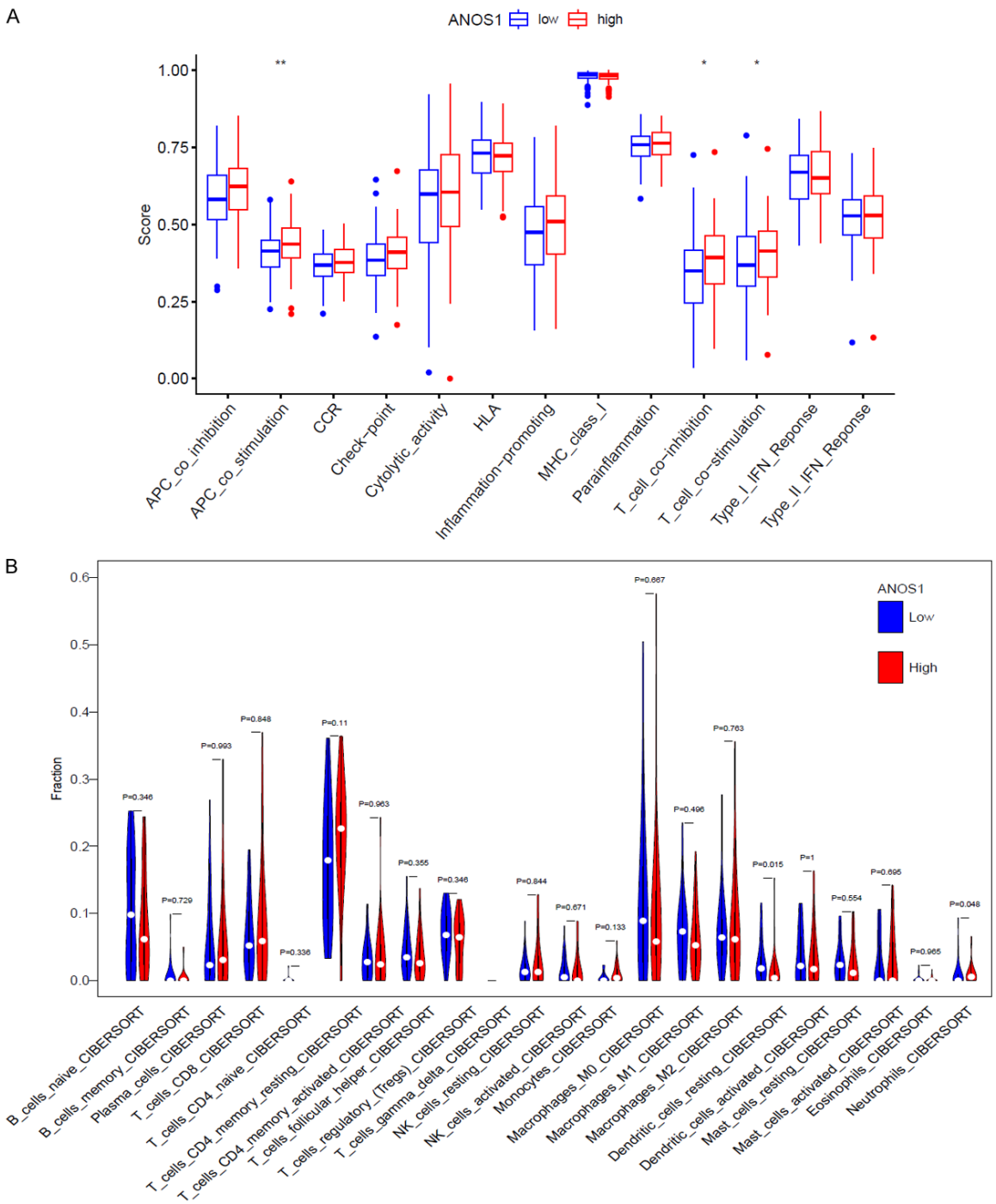
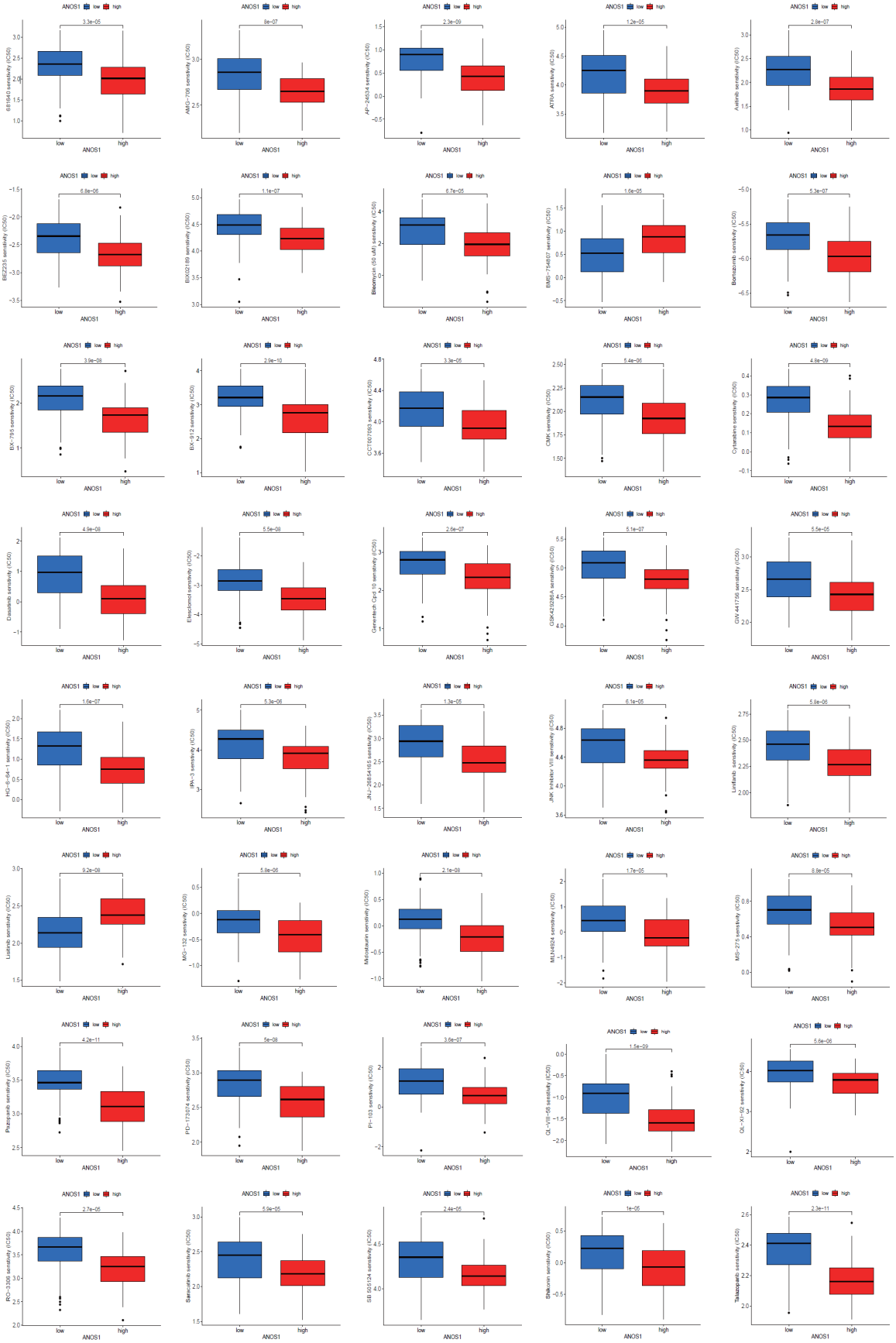


Figure S8. The correlation between the expression of *ANOS1* and immune cells as well as related immune pathways in the TCGA ESCA dataset. A. The comparison on the ssGSEA scores of 13 immune pathways by ANOS1 group (*P < 0.05, **P < 0.01, ***P < 0.001). B. The comparison on the immune cells according to CIBERSORT by *ANOS1* group.


ANOS1 low high
low
high
ANOS1

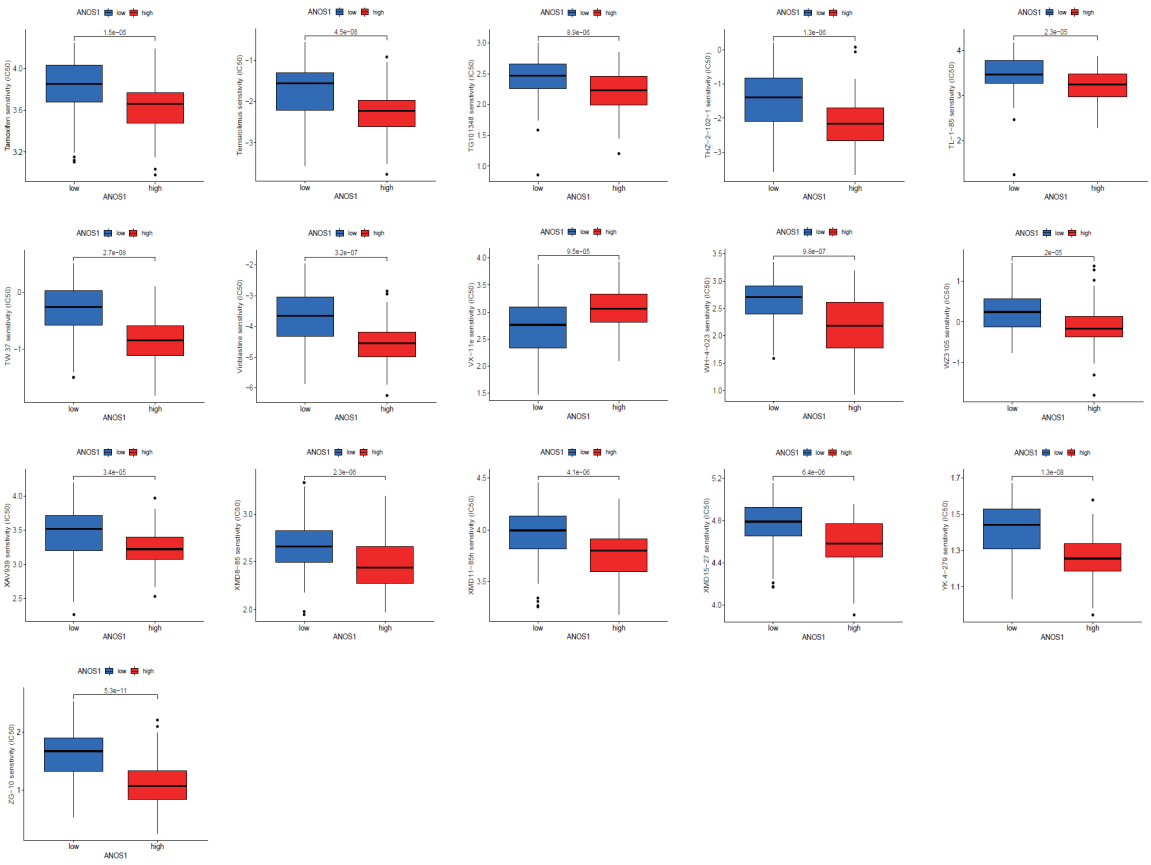


Figure S9. The comparison on chemosensitivity by the ANOS1 expression.

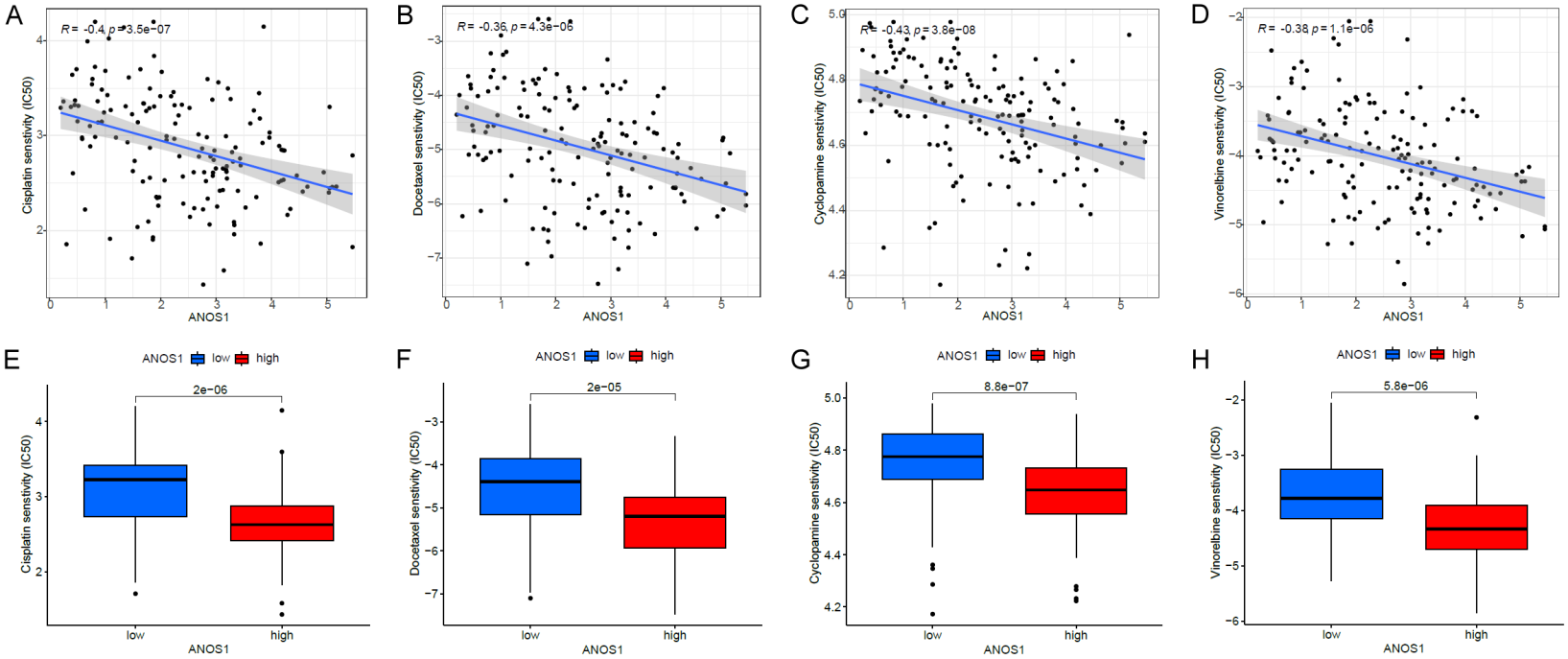


Figure S10. Potential chemotherapy drugs for esophageal cancer patients with high ANOS1 expression. The association between the ANOS1 expression and the estimated half inhibitory centration (IC50) for chemotherapeutics, including (A) cisplatin, (B) docetaxel, (C) cyclopamine, and (D) vinorelbine. The comparison between the ANOS1 group and the estimated IC50 for (E) cisplatin, (F) docetaxel, (G) cyclopamine, and (H) vinorelbine.