Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identification of Salivary Biomarkers from Gene Expression Data of Multiple Cancer Types Using Bioinformatics Analysis

Amirreza Safavinejad
Fasa University of Medical Sciences
Kiarash Zare
Shiraz University of Medical Sciences
Yaser Mansoori
Shiraz University of Medical Sciences
Elham Tayebi
Shiraz University of Medical Sciences
Mohammad Mehdi Naghizadeh
Mmnaghi2@yahoo.com
Fasa University of Medical Science

Research Article

Keywords: Biomarker, Cancer, Differentially expressed genes, Gene expression, Saliva

Posted Date: September 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10729148/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Identification of Salivary Biomarkers from Gene Expression Data of Multiple Cancer Types Using Bioinformatics Analysis

Amirreza Safavinejad[1], Kiarash Zare[2], Yaser Mansoori[3], Elham Tayebi[4], Mohammad Mehdi Naghizadeh[5 *]

[1]Systems Medicine Research Group, Student Research Committee, Fasa University of Medical Sciences, Fasa, Iran

[2]Shiraz Institute for Cancer Research, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran.

[3]Department of Medical Genetics, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran.

[4]Department of Oral and Maxillofacial Medicine, School of Dentistry, Shiraz University of Medical Sciences, Shiraz, Iran.

[5]Noncommunicable Diseases Research Center, Fasa University of Medical Science, Fasa, Iran

[*]E-mail: Mmnaghi2@yahoo.com, Address: Noncommunicable Disease Research Center, Fasa University of Medical Science, Fasa, Iran, Phone: +989177304178

Abstract

Objective: Cancer is recognized as the primary factor causing clinical and socio-economic pressure among diseases. This study investigates common salivary mRNA biomarkers in five types of cancer: ovarian, stomach, breast, lung, and pancreatic, using gene expression profiling by microarray.

Methods*:* The Gene Expression Omnibus (GEO) database was utilized to obtain salivary datasets and Differentially expressed genes (DEGs) were identified using GEO2R. Gene ontology and pathway analyses were conducted via the Enrichr database. External validation was performed using The Cancer Genome Atlas (TCGA) data, and the expression of immune subgroups was analyzed through the Tumor-Immune System Interaction Database (TISIDB).

Results: This study analyzed datasets including control and cancer groups and identified 80 common DEGs across the datasets. Genes such as IRAK3, SERPINA1, SUB1, and RBM6 showed both significant differences in expression in cancerous tissues and association with reduced survival rates. Additionally, gene expression correlated with immune cell infiltration in various cancers, suggesting that IRAK3 and RBM6 may serve as markers for immunotherapy efficacy.

Conclusion: This study identifies key differentially expressed genes associated with cancer prognosis and immune infiltration, and introduces IRAK3 and RBM6 as potential biomarkers for multi-cancer screening, risk stratification and as companion biomarkers for therapy selection.

Keywords: Biomarker, Cancer, Differentially expressed genes , Gene expression, Saliva

Introduction

In general, cancer represents the greatest clinical, social, and economic impact in terms of cause-specific Disability-Adjusted Life Years (DALYs) compared to all other human diseases [1]. Cancer ranks as the second most common cause of death in the United States overall, and it is the leading cause of death for individuals under 85 years of age. A study projects that in 2024, approximately 611,720 people in the U.S. will succumb to cancer, which equates to about 1,680 deaths each day [2]. Nowadays, the use of imaging methods, sampling, and laboratory data based on the sampling are common methods in cancer diagnosis [3,4].

Cancer diagnosis is one of the challenges in the cancer studies. It has primarily relied on various imaging modalities, including computed tomography (CT) and magnetic resonance imaging (MRI), and invasive tissue sampling approaches like bronchoscopy and needle biopsy techniques [4-7]. The limitations of these methods make it difficult to achieve some goals of the health care system, such as maximum screening of patients; For example, many of these diagnostic processes require the co-working between different of specialists such as radiology and pathology to be completed [8]. Another obstacle, is the invasiveness of some of these methods, which makes the patient less willing to participate in the diagnostic process [9]. The issue of the economic cost of cancer screening tests is also one of the effective factors in people's participation in timely cancer screening, which alongside its effect on the whole health care system's costs, it leads us to achieve more cost-effective methods [10]. All these factors imply a need to move toward more cost-effective, less invasive and more accessable methods for cancer screening and diagnosis.

Molecular biomarkers for cancer diagnosis were more enhanced as new high-throughput technologies like microarrays and sequencing developed. One of the important technologies that has caused new developments in the field of cancer diagnosis and risk assessment is the RNA microarray technique, which is used to show and describe thousands of gene expressions and generate gene expression data in the form of microarray [11]. Furthermore, the development of biochemical biomarkers, such as proteins, DNAs, mRNAs, and microRNAs, has opened up new methods for the diagnosis and treatment of cancer [12]. Although molecular biomarkers are an accurate and fast diagnostic tool, they must access the target tissues and have invasive sampling processes.

Saliva is an efficiently accessible, acidic biological fluid consisting of secretions from the three major salivary glands and minor glands, gingival crevicular fluid, cell debris, plaque, bacteria, nasal and bronchial secretions, lining cells, blood, and exogenous substances [13,14]. Saliva is used as a diagnostic aid in an increasing number of clinical situations and in systemic diseases that can affect the function and composition of the salivary glands, such as Mucoepidermoid Carcinoma, cystic fibrosis, and adrenocortical diseases [15,16]. It has many diagnostic advantages, such as non-invasiveness, cost-effectiveness, the possibility of being used as a screening tool in large populations, and the diagnosis of systemic diseases [13]. The term "salivaomics" was introduced to highlight the rapid development of knowledge about various "omics" constituents of saliva [17]. Extracting these constituents from saliva and using them as biomarkers has been one of the subjects focused on by researchers in order to create and use simple and

low-cost saliva tests as diagnostic tests in various cancers [18]. One of the salivaomics markers is mRNA, that has been repeatedly studied as a research target in cancer diagnosis. Messenger RNA is the direct precursor of proteins and in general the corresponding levels are correlated in cells and tissue samples [19]. As it was discovered that human messenger RNAs are present in cell-free form in saliva, mRNAs were profiled on gene expression arrays and divided into groups of "control" and "cancerous" [19,20].

In this study, according to the data collected in salivary microarray datasets related to cancer, we examined the levels of salivary mRNAs in cancer patients and identified common biomarkers in five types of cancer: ovarian, gastric, breast, lung, and pancreatic (The workflow of the study has been shown in the figure 1).

[Insert **Figure 1. Workflow of study**]

Materials and methods

Datasets

The Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) was used to obtain the dataset [21]. A comprehensive online search was conducted for the keywords "cancer", "saliva" and "salivary". We selected the datasets that contain the cancer and normal sample data. Only datasets from the Affymetrix Human Genome U133 Plus 2.0 Array (GPL570) platform were selectively included in this study due to their compatibility with GEO2R analysis.

Differentially expressed genes

We used GEO2R tools to identify differentially expressed genes (DEGs). GEO2R is an online tool based on the "limma" package in R programming. We analyzed the normal group compared with the cancer group to calculate the DEGs and defined a formula for the identified top 3000 probes in each dataset.

The formula integrates both the statistical significance p-value and the Logarithm (Fold change) into a single value [22]. Larger values indicate stronger biological relevance (high effect size + high significance):

$$-\log (\text{p-value}) * |\log (\text{fold change})|$$

Then, Venn diagram online tools (https://www.interactivenn.net/) were used to find common DEGs between different cancer datasets [23].

Gene ontology and enrichment pathway analysis

GO functional analysis and KEGG pathway analysis were carried out to predict the potential functions of the DEGs by using the Enrichr database (https://maayanlab.cloud/Enrichr/). The top 20 items of the biological process (BP), molecular function (MF), cellular component (CC) categories, and KEGG pathways were then sorted and presented in figures [24]. FDR -adjusted $P < 0.05$ was considered statistically significant.

External validation and survival analysis

We examined DEGs in RNA-seq datasets associated with breast invasive carcinoma (BRCA), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), ovarian serous cystadenocarcinoma (OV), and pancreatic adenocarcinoma (PAAD) stomach adenocarcinoma (STAD), which are available in the TCGA (The Cancer Genome Atlas) database. GEPIA2 is an online tool based on the TCGA database, which is utilized for gene expression profiles of human cancer and normal tissues.

GEPIA2 (http://gepia2.cancer-pku.cn/) was used to validate whether identified DEGs are significantly expressed in the cancer groups compared to the normal groups, confirming their differential expression patterns [25]. Also, survival analysis was used to analyze the prognostic value of the common genes. The Kaplan Meier-plotter dataset (https://kmplot.com/analysis/) is an online tool for survival analysis that was used to analyze the prognostic value of the stage-specific genes [26].

Immunity analysis

The TISIDB database (http://cis.hku.hk/TISIDB/index.php) combines various data types to study gene interactions in different immune tumors [27]. The relationship between gene expression and immune infiltration levels (lymphocyte, immunoinhibitor, immunostimulator, MHC molecules, chemokine and receptors) were explored through the TISIDB database. Statistical assessment was performed using a Spearman correlation to screen the appropriate candidates. (FDR adjusted $P < 0.05$ as a cut-off value)

After extracting the immune subtype expression data from the TISIDB database, these subtypes were analyzed to determine the significantly expressed immune subtypes. These six immune subtypes (C1-C6) represent distinct features of the tumor microenvironment (TME) that cut across traditional cancer classifications [28]. C1 (Wound Healing) had elevated expression of angiogenic genes, a high proliferation rate, and a Th2 cell bias to the adaptive immune infiltrate. C2 (IFN-γ dominant) had the highest M1/M2 macrophage polarization, a strong CD8 signal, and, together with C6, the greatest TCR diversity. C2 also showed a high proliferation rate, which may override an evolving Type I immune response. C3 (inflammatory) was defined by elevated Th17 and Th1 genes, low to moderate tumor cell proliferation, and, along with C5, lower levels of aneuploidy and overall somatic copy number alterations than the other subtypes. C4 (lymphocyte depleted) displayed a more prominent macrophage signature, with Th1 suppressed and a high M2 response. C5 (Immunologically Quiet) exhibited the lowest lymphocyte and highest macrophage responses, dominated by M2 macrophages. IDH mutations were enriched in C5 over C4, suggesting an association of IDH mutations with favorable immune composition. Indeed, IDH mutations associate with TME composition and decrease leukocyte chemotaxis, leading to fewer tumor-associated immune cells and better outcomes. C6 (TGF-β dominant) displayed the highest TGF-β signature and a high lymphocytic infiltrate with an even distribution of Type I and Type II T cells. These six categories represent features of the TME that largely cut across traditional cancer classifications to create groupings and suggest certain treatment approaches may be independent of histologic type.

Results

Datasets

Datasets were selected for analysis (Table 1) that had control and cancer groups. Among the datasets, those that included control groups and were based on the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array) were selectively included in this study. Datasets related to our search were available, including GSE14245, GSE20266, GSE29220, GSE64951, and GSE32175.

Table 1. Datasets were selected for finding salivary biomarkers

GEO accession	Cancer	Platform	Samples		Articles PMID
GSE14245	Pancreatic cancer	GPL570[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	24	Cancer (n) = 12 Normal (n) = 12	19931263
GSE20266	Breast cancer	GPL570[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	20	Cancer (n) = 10 Normal (n) = 10	21217834
GSE29220	Ovarian cancer	GPL570[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	22	Cancer (n) = 11 Normal (n) = 11	22095100
GSE64951	Gastric cancer	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	94	Cancer (n) = 63 Normal (n) = 31	30097497
GSE32175	Lung cancer	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	20	Cancer (n) = 10 Normal (n) = 10	22689099

Differentially expressed genes

GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) was used to pick out the DEGs in tissues. Then the first 3000 DEGs based on the formula with the highest fold change and lowest significance of each dataset were selected for further analysis. As shown in figure 2, 80 DEGs were common in all datasets (supplementary file 1).

[Insert **Figure 2. Venn diagram results of datasets.** Genes with differential expression that were overlapping in four or five cancers were selected for further analysis (A dashed line is drawn around them).]

Gene ontology and enrichment pathway analysis

Gene ontology (GO) and KEGG pathway enrichment analyses of overlapping DEGs were performed through Enrichr. According to our analysis, for biological process the top 10 exhibiting the most significant differences include positive regulation of biosynthetic process (GO:0009891, *P*-value: 7.25E-4), positive regulation of myelination (GO:0031643, *P*-value: 8.49E-4), myelination in peripheral nervous system (GO:0022011, *P*-value: 0.001), negative regulation of transmembrane receptor protein Serine/Threonine Kinase signaling pathway (GO:0090101, *P*-value: 0.001), regulation of transforming growth factor beta receptor signaling pathway (GO:0017015, *P*-value: 0.001), positive regulation of MAPK cascade (GO:0043410, *P*-value: 0.002), myeloid cell differentiation (GO:0030099, *P*-value: 0.002),

regulation of apoptotic process (GO:0042981, *P*-value: 0.002), positive regulation of extracellular matrix organization (GO:1903055, *P*-value: 0.002) and, positive regulation of nervous system process (GO:0031646, *P*-value: 0.002) (Figure3). The top 10 of molecular functions exhibiting the most significant differences including NAD-retinol dehydrogenase activity (GO:0004745, *P*-value: 0.003), RNA binding (GO:0003723, *P*-value: 0.01), transcription regulatory region nucleic acid binding (GO:0001067, *P*-value: 0.01), protease binding (GO:0002020, *P*-value: 0.01), serine-type endopeptidase inhibitor activity (GO:0004867, *P*-value: 0.02), DNA Binding (GO:0003677, *P*-value: 0.02), 4-galactosyl-N-acetylglucosaminide 3-alpha-L-fucosyltransferase activity (GO:0017083, *P*-value: 0.02), ADP-D-ribose modification-dependent protein binding (GO:0160002, *P*-value: 0.02), calcium-dependent protein serine/threonine phosphatase activity (GO:0004723, *P*-value: 0.02) and, nucleosome binding (GO:0031491, *P*-value: 0.03) (Figure3). The top KEGG pathways exhibiting the most significant differences including axon guidance (P-value: 0.006), rheumatoid arthritis (P-value: 0.006), NF-kappa B signaling pathway (P-value: 0.008), Rap1 signaling pathway (P-value: 0.01), TNF signaling pathway (P-value: 0.01), autophagy (P-value: 0.02), notch signaling pathway (P-value: 0.02), MAPK signaling pathway (P-value: 0.03), retinol metabolism (P-value: 0.03), adipocytokine signaling pathway (P-value: 0.03), adherens junction (P-value: 0.03), PPAR signaling pathway (P-value: 0.04) microRNAs in cancer (P-value: 0.04), ECM-receptor interaction (P-value: 0.05), PD-L1 expression and PD-1 checkpoint pathway in cancer (P-value: 0.05), viral protein interaction with cytokine and cytokine receptor (P-value: 0.06), and Ras signaling pathway (P-value: 0.07) (Figure3). Other results of GO and KEGG pathway enrichment analyses were presented in the supplementary file 2.

[Insert **Figure 3. Gene ontology and enrichment pathway analysis.** The most significant terms of molecular function, biological process and KEGG pathway for overlapping DEGs were identified using Enrichr. The terms with a p-value <0.05 were considered significant.]

External validation and survival analysis

Out of a total of 81 selected genes, IRAK3, SERPINA1, SUB1, RBM6, and TNXB had significant expression differences in external evaluation. Based on log (fold change), the manner of gene expression regulation was investigated in all cancers (Figure 4).

[Insert **Figure 4. Expression regulation of genes in lung, ovarian, gastric, breast and pancreatic cancers.**]

The results of this study were validated in original tissues using the GEPIA2 database. Based on results, IRAK3, SERPINA1, TNXB, SUB1, and RBM6 have significant differential expression in cancer tissues and have meaningful prognostic value. However, we were able to identify these DEGs in saliva. Expression of TNXB was significantly downregulated in the BRCA, LUAD, OV, and STAD cancers and significantly upregulated in PAAD cancer compared to their normal groups (Figure 5). Expression of RBM6 was significantly downregulated in the BRCA, LUAD, OV, and STAD cancers compared with the normal group. But no significant difference was detected in PAAD compared with the normal group (Figure 5). Expression of SUB1 was significantly upregulated in OV and STAD cancer compared with the normal group. But no significant difference was detected in BRCA and LUAD compared with the normal group (Figure 5). Expression boxplot of SERPINA1 was significantly upregulated in the BRCA, OV, PAAD, and STAD cancers compared with the normal group. But no significant difference was detected in LUAD cancer compared with the normal group (Figure 5). Expression of IRAK3 was significantly downregulated in the BRCA, LUAD, and OV and significantly upregulated in the PAAD cancer group compared with the normal group. But no significant difference was detected in STAD cancer compared with the normal group (Figure 5).

Survival analysis by GEPIA2 indicated that the expression of SUB1, SERPINA1, RBM6, and IRAK3 was significantly related to the overall reduction in survival rate of patients with cancer. On the other hand, although the expression of TNXB was related to the reduction of lifespan, it was not significantly related to the prognosis of patients (Figure 6).

[Insert **Figure 5. mRNA expression of identified overlapping DEGs.** Comparison of expression levels of TNXB, BMR6, SUB1, SERPINA1, and IRAK3 in breast cancer (BRCA), lung cancer (LUAD), ovarian cancer (OV), pancreatic cancer (PAAD), and gastric cancer (STAD) with the normal groups using the GEPIA2 database. P-value < 0.05.]

[Insert **Figure 6. Kaplan–Meier overall survival analysis for overlapping DEGs.** The expression of SUB1, SERPINA1, RBM6, and IRAK3 was significantly related to the overall survival of patients with cancer.]

Immunity analysis

The role of gene expression on immune subtypes among five cancers was explored with the TISIDB website and presented in figure 7. IRAK3, RBM6, SERPINA1, and SUB1 expression was related to different immune cell infiltration in breast, gastric, ovarian, pancreatic, and lung cancer. The heat map demonstrates that IRAK3 shows a generally positive correlation with immune infiltrations, including lymphocytes, immune inhibitors, MHC molecules, and chemokines across all cancers. A predominantly negative correlation for RBM6 was shown with almost all of the immune infiltrations. There is a chance for IRAK3 and RBM6 to become biomarkers to predict the efficacy of immunotherapy.

Analysis of immune subtype expression data revealed that the Inflammatory subtype (C3) and TGF-β dominant subtype (C6) exhibited the highest significant expression levels (Figure 8). Notably, RBM6 and IRAK3 expressions were most strongly associated with subtypes C3 and C6 across all cancer types. Additionally, SERPINA1 expression showed the highest association with subtype C6, while SUB1 expression was most closely associated with IFN-γ dominant subtype (C2).

[Insert **Figure 7. Correlation of immune infiltration expression and SUB1, IRAK3, SERPINA1 and RBM6.** The immune components including lymphocytes, immune-inhibitors, immune-stimulators, MHC molecules, chemokine, and receptors were found in Breast invasive carcinoma(B), Lung adenocarcinoma(L), Ovarian serous cystadenocarcinoma(O), Pancreatic adenocarcinoma(P), Stomach adenocarcinoma(S).]

[Insert **Figure 8**. **Immune subtype in breast invasive carcinoma (BRCA), lung adenocarcinoma (LUAD), ovarian serous cystadenocarcinoma (OV), pancreatic adenocarcinoma (PAAD) and, stomach adenocarcinoma (STAD)**]

Discussion

Saliva, as an available fluid specimen, contains diverse molecules, including DNA/RNA, proteins, metabolites, and microbiota, which reflect tumor-specific changes and systemic conditions. Salivary biomarkers have emerged as a promising non-invasive tool for cancer screening and detection, offering advantages in early diagnosis and treatment monitoring. In this study, common differentially expressed genes in breast, pancreatic, lung, stomach, and ovarian cancers were identified in the saliva samples of cancer groups compared to the control. The results showed that the TNXB, SUB1, SERPINA1, RBM6, and IRAK3 genes have significant expression differences in five cancers, and SUB1, SERPINA1, RBM6, and IRAK3 have significant prognostic values. The correlation between the immune system and SUB1, SERPINA1, RBM6, and IRAK3 genes was investigated, and the results show a high and significant correlation with lymphocytes, immune inhibitors, immune stimulators, MHC molecules, chemokines, and receptors.

In our study, TNX had differential expression in salivary samples, and evaluation based on TCGA data also showed that it had significant differential expression in breast, lung, pancreatic, ovarian, and gastric cancer tissues. TNXB encodes the extracellular matrix protein tenascin-X, has emerged as a significant factor in the type and stage of the cancer [29,30]. In some different cancers, such as breast, lung, colorectal, prostate, stomach, and liver cancers, TNXB expression is downregulated during progression and associated advanced stages, indicating a tumor-suppressive function in cancer mechanisms [31,32]. Additionally, TNXB has been involved in modulating tumor microenvironment interactions, including immune response, and cancer-associated fibroblast activation, therapy response and impacting prognosis in cancers like head and neck squamous cell carcinoma [33]. According to previous studies and the results of this study, it seems that the identification of TNX in saliva can be considered for diagnostic and therapeutic purposes.

In our study, the IRAK3 gene was expressed differentially in lung, ovarian, breast, and gastric cancer in comparison to the control group. IRAK3 regulates TLR signaling, promoting immune homeostasis but suppressing anti-tumor immunity, with high expression in tumors indicating anti-inflammatory pathways and poorer immune checkpoint responses [34,35]. In lung cancer,

IRAK3 plays a crucial role in tumorigenesis, progression, immune evasion, patient prognosis, and therapy responsiveness in lung cancer, potentially acting as a prognostic predictor and immunotherapeutic indicator [36,37]. Caba et al. suggested IRAK3 as an accurate biomarker for diagnosis in pancreatic cancer by gene expression profiling of blood [38]. Zheng et al. found that IRAK3 plays a role in metastasis in gastric cancer [39]. Also, some studies show it has a considerable function in colorectal cancer [40], glioma [41], and melanoma [42]. Furthermore, the survival analysis based on the GEPIA database demonstrated significant prognostic value in lung, ovarian, breast, gastric, and panceartic cancer tissue. While evidence shows that IRAK3 plays an important role in cancerous tissues, our study showed that it also has a significant difference in expression in saliva samples, which can be considered for diagnostic and therapeutic purposes.

Another gene with differential expression was RBM6, which the results of this study showed was significantly different in the saliva of lung, breast, and pancreatic cancers compared to the control group. The loss of RBM6 expression has been associated with increased tumor growth and progression, suggesting its critical role in maintaining genomic integrity and regulating cell proliferation [43]. As an alternative splicing factor, RBM6 regulates the splicing of various pre-mRNAs involved in critical cellular processes, including those linked to tumorigenesis. By modulating the splicing patterns of specific genes, RBM6 can influence the expression of proteins that either promote or inhibit cancer progression [44]. Some studies have pointed out this double functionality of the gene in the tumor environment. Liu et al. studied the role of RBM6's effect on the migration of prostate tumor cells and concluded that its inhibitory effect is only present when there is a high expression of the ZEB1 gene, and in a normal condition, it would promote the tumor cell migration [45]. Along side that, tumor-suppressing function of RBM6 has been studied for many cancers, including hepatocellular carcinoma [46] and laryngocarcinoma [47]. On the other hand, Duan et al. has nominated RBM6 as a diagnostic serum biomarker for pancreatic cancer, as it has shown a higher expression in pancreatic cancer patients compared to paracancerous tissue [48]. In this study, a negative correlation with almost all of the immune infiltrations was obtained, and the survival analysis pointed to its relation to the overall reduction in survival rate of patients with cancer.

Serpin family A member 1 (SERPINA1) is the gene responsible for encoding the protein alpha1-antitrypsin (A1AT) that is the dominant protease inhibitor of the cellular plasma [49]. But its biological functions extend beyond protease inhibition; it is involved in modulating inflammatory responses and tissue repair mechanisms [50]. Different studies have investigated the role of SERPINA1 in mechanisms and progression of multiple cancers, including gastric cancer [51,52], breast cancer [53,54], and colorectal cancer [55,56]. Kuai et al. has done a detailed pan-cancer bioinformatic analysis on SERPINA1 using TCGA datasets, including expression level analysis, immune analysis, and survival analysis, which resulted in an emphasis on SERPINA1's expression level as an indicator for diagnostic and prognostic assessment of different cancers [49]. In this current study, the analysis showed the upregulation of SERPINA1 in the salivary samples of patients with pancreatic, gastric, and breast cancer and downregulation in ovarian cancer compared to the normal groups. The external validation showed its upregulation in all four types of cancer, along with a significant prognostic value in the survival analysis. The immune analysis demonstrated a positive correlation between SERPINA1 and immune infiltrations in ovarian cancer patients.

SUB1 (PC4) has been shown to regulate the expression of several oncogenes, including PLK1 (polo-like kinase 1), BUB1B (BUB1 mitotic checkpoint serine/threonine kinase), and c-MYC. These interactions suggest that SUB1 acts as a critical mediator between upstream signaling pathways and the transcriptional machinery that drives oncogenic gene expression [57,58]. Its function in colorectal cancer metastasis [59], prostate cancer [57], non-small cell lung cancer [60], astrocytoma [61], and prognosis of liver cancer patients [62] has been studied and signified. The evidence on SUB1 in breast cancer seems to be contradictory. Although there are functional studies providing experimental support for the oncogenic role of SUB1 [63], Sikder et al. showed that the expression analysis across thirteen breast cancer cell lines with varying molecular subtypes and invasive potential revealed substantial downregulation of PC4 in three highly aggressive cell lines (ZR-75-1, HCC38, HCC1806) compared to non-tumorigenic MCF10A cells [64]. The current study showed that the expression of SUB1 is upregulated in the salivary samples of lung and pancreatic cancer and downregulated in the samples of ovarian and gastric cancer patients compared to the normal group. The external validation using the GEPIA2 database demonstrated a significant upregulation in gastric, pancreatic, and ovarian cancer patients, and its prognostic value was considerable. The immune analysis showed a significant negative correlation between SUB1 expression and immune infiltrations in breast cancer.

This study identified salivary biomarkers in gastric, breast, lung, ovarian, and pancreatic cancers. External evaluation showed that SUB1, SERPINA1, RBM6, and IRAK3 genes have high expression and prognostic values. The results of this study can be used for diagnostic and therapeutic purposes. Considering that the genes studied in this article significantly correlated with the immune system, they can be candidates for monitoring patients in follow-up. If further studies can confirm and evaluate these results, it may be possible to use immuno-biomarkers for cancer screenings.

The general idea of this study is that by considering the human body as a complex system, we expect to see the emergence of interactions between different components of tumor pathogenesis in other parts of the body. The conductors of this study were up to validate this hypothesis that the indirect results of tumor biological component interactions can be detected so significantly in saliva that they can be proposed as a basis for future studies in the field of cancer biology, including detection, risk stratification, and therapy candidation.

The datasets used in this study used a common platform for their sample gathering, which was an important factor for the quality of this study. Although the process of DEG candidation was done in salivary samples, the results were externally validated using the data available on the TCGA database which contains data from tissue samples, which means that the differentially expressed genes in saliva were also differentially expressed in cancer tissues in comparison to normal tissues.

Conclusion

This study has used bioinformatics analysis approaches to identify salivary biomarkers in gastric, breast, lung, ovarian, and pancreatic cancers. After external validation SUB1, SERPINA1, RBM6, and IRAK3 genes have been candidated as expression and prognostic markers. Although more experimental studies are required to validate our study's results, we hope our findings will assist in identifying salivary biomarkers for cancer screening.

Abreviations

A1AT	Alpha1-antitrypsin
BP	Biological Process
BRCA	Breast Invasive Carcinoma
BUB1B	BUB1 mitotic checkpoint serine/threonine kinase
CC	Cellular Component
CT	Computed Tomography
DALY	Disability-Adjusted Life Years
DNA	Deoxyribonucleic Acid
DEG	Differentially Express Genes
ECM	Extracellular matrix
FDR	False Discovery Rate
GEO	Gene Expression Omnibus
IRAK3	Interleukin 1 Receptor Associated Kinase 3
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MAPK	Mitogen-activated protein kinases
MF	Molecular Function
MHC	Major histocompatibility complex
MRI	Magnetic Resonance Imaging
MRNA	Messenger Ribonucleic Acid
MYC	Myelocytoma
OV	Ovarian cancer
PAAD	Pancreatic adenocarcinoma
PLK1	Polo-like kinase 1
PPAR	Peroxisome Proliferator-Activated Receptor
RBM6	RNA Binding Motif Protein 6
RNA	Ribonucleic Acid
SERPINA1	Serpin Family A Member 1
STAD	Stomach adenocarcinoma

SUB1	SUB1 Regulator Of Transcription
TCGA	The Cancer Genome Atlas
TISIDB	Tumor-Immune System Interaction Database
TLR	Toll-like receptors
TME	Tumor microenvironment
TNF	Tumor Necrosis Factor
TNXB	Tenascin XB
ZEB1	Zinc finger E-box Binding homeobox 1

Authors contribution

Study conceptualization and design: A.S and K.Z .Data collection: A.S and E.T. Data analysis: A.S and K.Z. Manuscript writing: A.S, K.Z, Y.M, E.T and M.M.N. Interpretation of results: K.Z, Y.M and M.M.N. Project supervising: M.M.N. All authors read and approved the submitted version of the article.

Acknowledgments

No comments.

Statements and Declarations

Ethical considerations

This project was approved and supervised by the Ethics Committee of Fasa University of Medical Sciences. (Code: IR.FUMS.REC.1403.025).

Consent to participate

Not applicable.

Consent for publication

Not applicable

Declaration of conflicting interest

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

Funding statement

This study was not funded by any organization.

Declaration of Generative AI in the Writing Process

During the preparation of this work, the authors used Grammarly Premium's advanced features to improve readability and refine sentence structure. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Data availability

The datasets generated and/or analyzed during the current study are available in the GEO repository with accession numbers: GSE14245, GSE20266, GSE29220, GSE64951, and GSE32175. It is a free, publicly accessible repository that stores a large number of gene function and expression data. The details can be found in Table 1. In this project, we also used some online databases, resources, and tools such as the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo), Venn diagram online tools (https://www.interactivenn.net/), Enrichr database (https://maayanlab.cloud/Enrichr/), TCGA (The Cancer Genome Atlas) database. (https://portal.gdc.cancer.gov/), GEPIA2

(http://gepia2.cancer-pku.cn/), Kaplan-Meier Plotter dataset (https://kmplot.com/analysis/), TISIDB database (http://cis.hku.hk/TISIDB/index.php). The Methods section explains the details and parameters used.

References

1 Mattiuzzi, C. & Lippi, G. Current Cancer Epidemiology. *J Epidemiol Glob Health* **9**, 217-222 (2019). https://doi.org:10.2991/jegh.k.191008.001

2 Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA: a cancer journal for clinicians* **74**, 12-49 (2024).

3 Jafari, S. H. *et al.* Breast cancer diagnosis: Imaging techniques and biochemical markers. *Journal of cellular physiology* **233**, 5200-5213 (2018).

4 Ofiara, L. M., Navasakulpong, A., Beaudoin, S. & Gonzalez, A. V. Optimizing tissue sampling for the diagnosis, subtyping, and molecular analysis of lung cancer. *Frontiers in oncology* **4**, 253 (2014).

5 Mazzone, P., Jain, P., Arroliga, A. C. & Matthay, R. A. Bronchoscopy and needle biopsy techniques for diagnosis and staging of lung cancer. *Clinics in Chest Medicine* **23**, 137-158 (2002). https://doi.org:https://doi.org/10.1016/S0272-5231(03)00065-0

6 Zampa, V., Roselli, G. & Beltrami, G. MRI of bone tumors: advances in diagnosis and treatment assessment. *Imaging in Medicine* **2**, 325 (2010).

7 Trinh, T. W. *et al.* Bladder cancer diagnosis with CT urography: test characteristics and reasons for false-positive and false-negative results. *Abdominal Radiology* **43**, 663-671 (2018). https://doi.org:10.1007/s00261-017-1249-6

8 Bogot, N. R. & Shaham, D. SEMI-INVASIVE AND INVASIVE PROCEDURES FOR THE DIAGNOSIS AND STAGING OF LUNG CANCER II: Bronchoscopic and Surgical Procedures. *Radiologic Clinics of North America* **38**, 535-544 (2000). https://doi.org:https://doi.org/10.1016/S0033-8389(05)70183-4

9 Cooper, C. P. & Gelb, C. A. Opportunities to expand colorectal cancer screening participation. *Journal of women's health* **25**, 990-995 (2016).

10 McCaul, K. D. & Tulloch, H. E. Cancer Screening Decisions. *JNCI Monographs* **1999**, 52-58 (1999). https://doi.org:10.1093/oxfordjournals.jncimonographs.a024209

11 Osama, S., Shaban, H. & Ali, A. A. Gene reduction and machine learning algorithms for cancer classification based on microarray gene expression data: A comprehensive review. *Expert Systems with Applications* **213**, 118946 (2023). https://doi.org:https://doi.org/10.1016/j.eswa.2022.118946

12 Seyed Hamed Jafari, Z. S., Arash Salmaninejad, Fatemeh Momeni, Mojgan Mokhtari, Javid Sadri Nahand, Majid Rahmati, Hamed Mirzaei, Mojtaba Kianmehr. Breast cancer diagnosis: Imaging techniques and biochemical markers. *Journal of Cellular Physiology* (2017).

13 Kaufman, E. & Lamster, I. B. The diagnostic applications of saliva—a review. *Critical Reviews in oral biology & medicine* **13**, 197-212 (2002).

14 Lee, Y.-H. & Wong, D. T. Saliva: an emerging biofluid for early detection of diseases. *American journal of dentistry* **22**, 241 (2009).

15 Mandel, I. D. The diagnostic uses of saliva. *Journal of Oral Pathology & Medicine* **19**, 119-125 (1990).

16 Lin, H. H., Limesand, K. H. & Ann, D. K. Current State of Knowledge on Salivary Gland Cancers. *Crit Rev Oncog* **23**, 139-151 (2018). https://doi.org:10.1615/CritRevOncog.2018027598

17 Nonaka, T. & Wong, D. T. Saliva diagnostics: salivaomics, saliva exosomics, and saliva liquid biopsy. *The Journal of the American Dental Association* **154**, 696-704 (2023).

18 Kaczor-Urbanowicz, K. E. *et al.* Saliva diagnostics–Current views and directions. *Experimental Biology and Medicine* **242**, 459-472 (2017).

19 Zimmermann, B. G. & Wong, D. T. Salivary mRNA targets for cancer diagnostics. *Oral Oncology* **44**, 425-429 (2008). https://doi.org:https://doi.org/10.1016/j.oraloncology.2007.09.009

20 John, M. A. S. *et al.* Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. *Archives of Otolaryngology–Head & Neck Surgery* **130**, 929-935 (2004).

21 Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research* **41**, D991-D995 (2012). https://doi.org:10.1093/nar/gks1193

22 Feng, J. *et al.* GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* **28**, 2782-2788 (2012). https://doi.org:10.1093/bioinformatics/bts515

23 Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* **16**, 169 (2015). https://doi.org:10.1186/s12859-015-0611-3

24 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97 (2016). https://doi.org:10.1093/nar/gkw377

25 Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* **47**, W556-w560 (2019). https://doi.org:10.1093/nar/gkz430

26 Győrffy, B. Integrated analysis of public datasets for the discovery and validation of survival-associated genes in solid tumors. *Innovation (Camb)* **5**, 100625 (2024). https://doi.org:10.1016/j.xinn.2024.100625

27 Ru, B. *et al.* TISIDB: an integrated repository portal for tumor–immune system interactions. *Bioinformatics* **35**, 4200-4202 (2019). https://doi.org:10.1093/bioinformatics/btz210

28 Thorsson, V. *et al.* The immune landscape of cancer. *Immunity* **48**, 812-830. e814 (2018).

29 Nakayama, K. *et al.* Tenascin XB Is a Novel Diagnostic Marker for Malignant Mesothelioma. *Anticancer Res* **39**, 627-633 (2019). https://doi.org:10.21873/anticanres.13156

30 Yuan, Y. *et al.* Tenascin-X is a novel diagnostic marker of malignant mesothelioma. *Am J Surg Pathol* **33**, 1673-1682 (2009). https://doi.org:10.1097/PAS.0b013e3181b6bde3

31 Liot, S. *et al.* Loss of Tenascin-X expression during tumor progression: A new pan-cancer marker. *Matrix Biol Plus* **6-7**, 100021 (2020). https://doi.org:10.1016/j.mbplus.2020.100021

32 Tucker, R. P. & Degen, M. Revisiting the Tenascins: Exploitable as Cancer Targets? *Front Oncol* **12**, 908247 (2022). https://doi.org:10.3389/fonc.2022.908247

33 Xu, Y., Li, J., Wang, J. & Deng, F. A novel CAF-cancer cell crosstalk-related gene prognostic index based on machine learning: prognostic significance and prediction of therapeutic response in head and neck squamous cell carcinoma. *J Transl Med* **22**, 645 (2024). https://doi.org:10.1186/s12967-024-05447-6

34 Tunalı, G. *et al.* IL-1 receptor-associated kinase-3 acts as an immune checkpoint in myeloid cells to limit cancer immunotherapy. *J Clin Invest* **133** (2023). https://doi.org:10.1172/jci161084

35 Rhyasen, G. W. & Starczynowski, D. T. IRAK signalling in cancer. *British Journal of Cancer* **112**, 232-237 (2015). https://doi.org:10.1038/bjc.2014.513

36 Zhou, Y. *et al.* IL-1 receptor-associated kinase 3 (IRAK3) in lung adenocarcinoma predicts prognosis and immunotherapy resistance: involvement of multiple inflammation-related pathways. *Translational Lung Cancer Research* **13**, 2139-2161 (2024).

37 Standiford, T. J. *et al.* TGF-β-induced IRAK-M expression in tumor-associated macrophages regulates lung tumor growth. *Oncogene* **30**, 2475-2484 (2011). https://doi.org:10.1038/onc.2010.619

38 Caba, O. *et al.* Transcriptional profiling of peripheral blood in pancreatic adenocarcinoma patients identifies diagnostic biomarkers. *Dig Dis Sci* **59**, 2714-2720 (2014). https://doi.org:10.1007/s10620-014-3291-3

39 Zheng, Y. *et al.* Multiple data sets to explore the key molecules and mechanism of lymph node metastasis in gastric cancer. *Discov Oncol* **16**, 606 (2025). https://doi.org:10.1007/s12672-025-02360-2

40 Kesselring, R. *et al.* IRAK-M Expression in Tumor Cells Supports Colorectal Cancer Progression through Reduction of Antimicrobial Defense and Stabilization of STAT3. *Cancer Cell* **29**, 684-696 (2016). https://doi.org:10.1016/j.ccell.2016.03.014

41 Wu, X., Ouyang, Y., Wang, B., Lin, J. & Bai, Y. Hypermethylation of the IRAK3-Activated MAPK Signaling Pathway to Promote the Development of Glioma. *Cancer Manag Res* **12**, 7043-7059 (2020). https://doi.org:10.2147/cmar.S252772

42 Geng, D. *et al.* Induction of IRAK-M in melanoma induces caspase-3 dependent apoptosis by reducing TRAF6 and calpastatin levels. *Commun Biol* **3**, 306 (2020). https://doi.org:10.1038/s42003-020-1033-y

43 Machour, F. E. *et al.* RBM6 splicing factor promotes homologous recombination repair of double-strand breaks and modulates sensitivity to chemotherapeutic drugs. *Nucleic Acids Research* **49**, 11708-11727 (2021). https://doi.org:10.1093/nar/gkab976

44 Li, Z. *et al.* The RNA-Binding Motif Protein Family in Cancer: Friend or Foe? *Frontiers in Oncology* **11** (2021). https://doi.org:10.3389/fonc.2021.757135

45 Liu, R. *et al.* The Oncopromoting Gene RBM6 Inhibits Prostate Tumour Cell Migration During Epithelial-to-Mesenchymal Transition. *Journal of Cellular and Molecular Medicine* **29**, e70397 (2025).

46 Ye, Z. *et al.* RNA-binding protein RBM6 acts as a tumor suppressor gene to inhibit the progression of hepatocellular carcinoma. *J buon* **26**, 402-408 (2021).

47 Wang, Q. *et al.* RNA-binding protein RBM6 as a tumor suppressor gene represses the growth and progression in laryngocarcinoma. *Gene* **697**, 26-34 (2019). https://doi.org:10.1016/j.gene.2019.02.025

48 Duan, B. *et al.* RNA-Binding Motif Protein 6 is a Candidate Serum Biomarker for Pancreatic Cancer. *Proteomics Clin Appl* **13**, e1900048 (2019). https://doi.org:10.1002/prca.201900048

49 Kuai, X., Lv, J., Zhang, J., Xu, M. & Ji, J. Serpin Family A Member 1 Is Prognostic and Involved in Immunological Regulation in Human Cancers. *International Journal of Molecular Sciences* **24**, 11566 (2023).

50 Ercetin, E. *et al.* Clinical significance of SERPINA1 gene and its encoded alpha1-antitrypsin protein in NSCLC. *Cancers* **11**, 1306 (2019).

51 Yang, J. *et al.* Identification of peptide regions of SERPINA1 and ENOSF1 and their protein expression as potential serum biomarkers for gastric cancer. *Tumor Biology* **36**, 5109-5118 (2015).

52 Kwon, C. *et al.* Serpin peptidase inhibitor clade A member 1 is a biomarker of poor prognosis in gastric cancer. *British journal of cancer* **111**, 1993-2002 (2014).

53 Chan, H. J. *et al.* SERPINA1 is a direct estrogen receptor target gene and a predictor of survival in breast cancer patients. *Oncotarget* **6**, 25815 (2015).

54 Yu, Q. *et al.* Exploration of SERPINA family functions and prognostic value in breast cancer based on transcriptome and in vitro analysis. *Environmental Toxicology* **39**, 1951-1967 (2024).

55 Ma, Y. *et al.* CEBPB-mediated upregulation of SERPINA1 promotes colorectal cancer progression by enhancing STAT3 signaling. *Cell Death Discovery* **10**, 219 (2024). https://doi.org:10.1038/s41420-024-01990-9

56 Li, L. *et al.* Plasma exosomal protein PLG and SERPINA1 in colorectal cancer diagnosis and coagulation abnormalities. *Journal of Cancer Research and Clinical Oncology* **149**, 8507-8519 (2023).

57 Chakravarthi, B. V. S. K. *et al.* MicroRNA-101 regulated transcriptional modulator SUB1 plays a role in prostate cancer. *Oncogene* **35**, 6330-6340 (2016). https://doi.org:10.1038/onc.2016.164

58 Chetverina, D., Vorobyeva, N. E., Gyorffy, B., Shtil, A. A. & Erokhin, M. Analyses of Genes Critical to Tumor Survival Reveal Potential 'Supertargets': Focus on Transcription. *Cancers (Basel)* **15** (2023). https://doi.org:10.3390/cancers15113042

59 Wang, H. *et al.* SUB1 promotes colorectal cancer metastasis by activating NF-κB signaling via UBR5-mediated ubiquitination of UBXN1. *Sci China Life Sci* **67**, 1199-1211 (2024). https://doi.org:10.1007/s11427-023-2429-5

60 Peng, Y. *et al.* Human positive coactivator 4 is a potential novel therapeutic target in non-small cell lung cancer. *Cancer Gene Ther* **19**, 690-696 (2012). https://doi.org:10.1038/cgt.2012.52

61 Chen, L. *et al.* Human positive coactivator 4 (PC4) is involved in the progression and prognosis of astrocytoma. *J Neurol Sci* **346**, 293-298 (2014). https://doi.org:10.1016/j.jns.2014.09.014

62 Yong-ping, H. *et al.* Prognostic value of SUB1 in liver cancer patients. *Lingnan Modern Clinics In Surgery* **20**, 425 (2020).

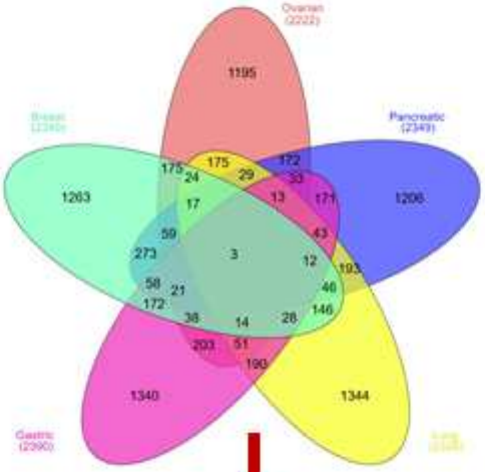
63 Luo, P. *et al.* Transcriptional positive cofactor 4 promotes breast cancer proliferation and metastasis through c-Myc mediated Warburg effect. *Cell Commun Signal* **17**, 36 (2019). https://doi.org:10.1186/s12964-019-0348-0

64 Sikder, S. *et al.* Chromatin protein PC4 is downregulated in breast cancer to promote disease progression: Implications of miR-29a. *Oncotarget* **10**, 6855-6869 (2019). https://doi.org:10.18632/oncotarget.27325

Gene Expression Omnibus



Cancers of Breast, Gastric, Ovarian, Pancreatic, Lung

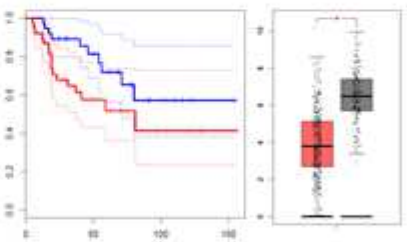


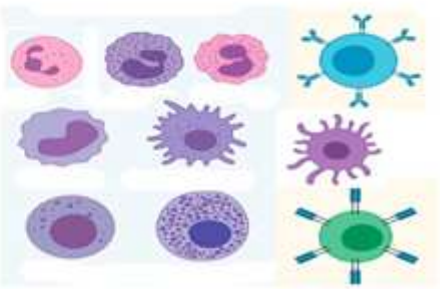
Enrichr

Saliva common DEGs

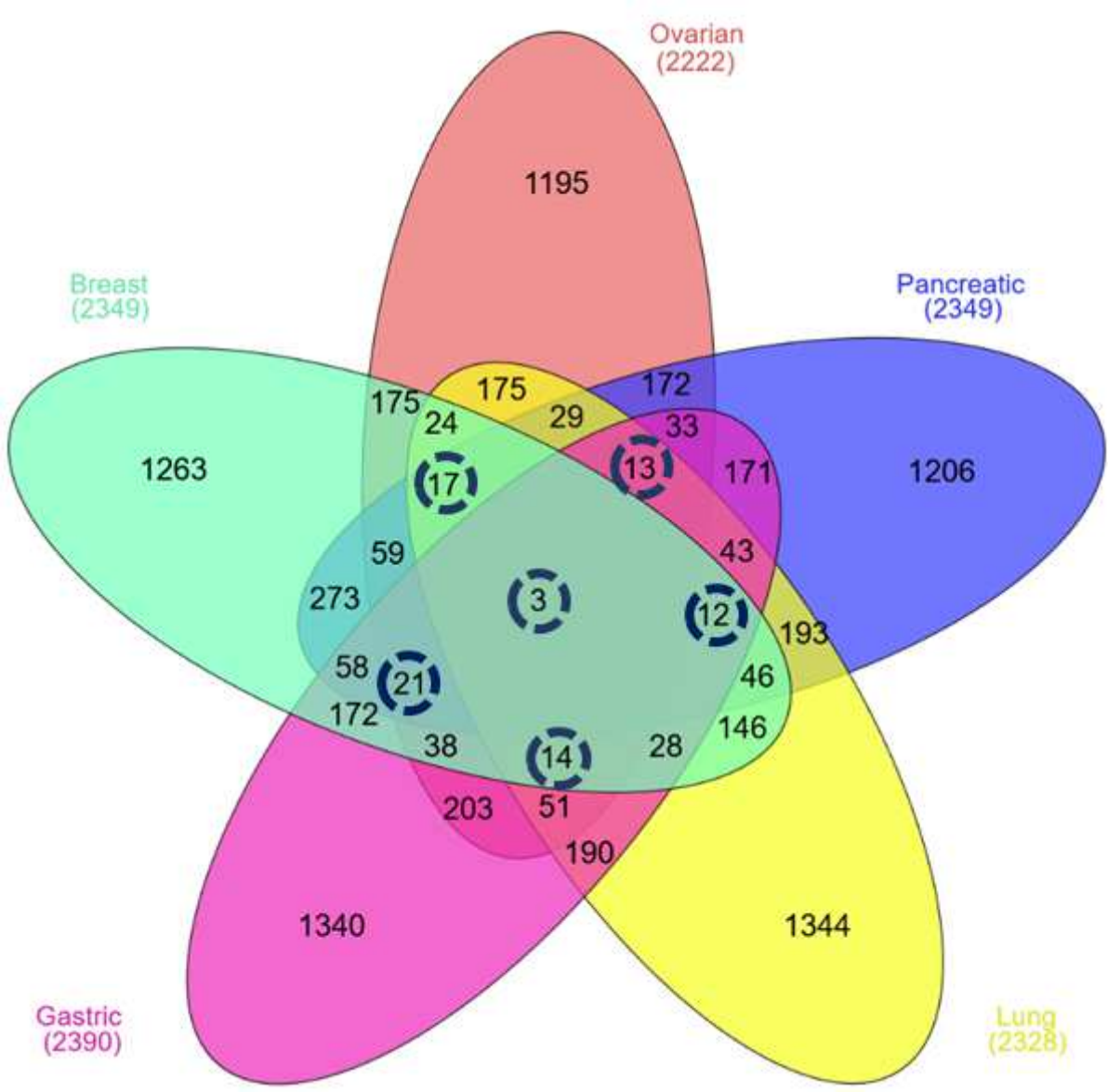
GEPIA2

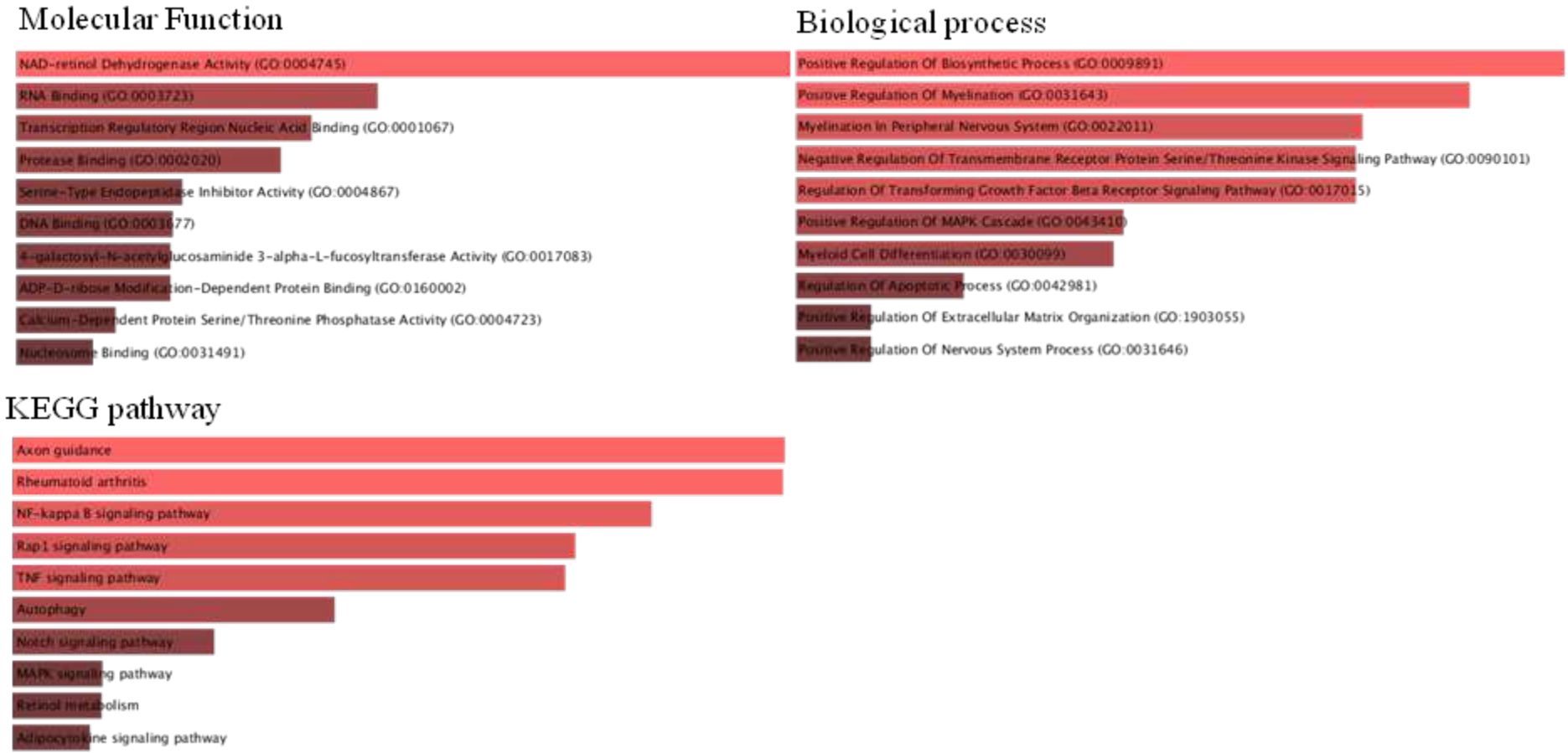
* Gene Ontology
* Enrichment Pathway





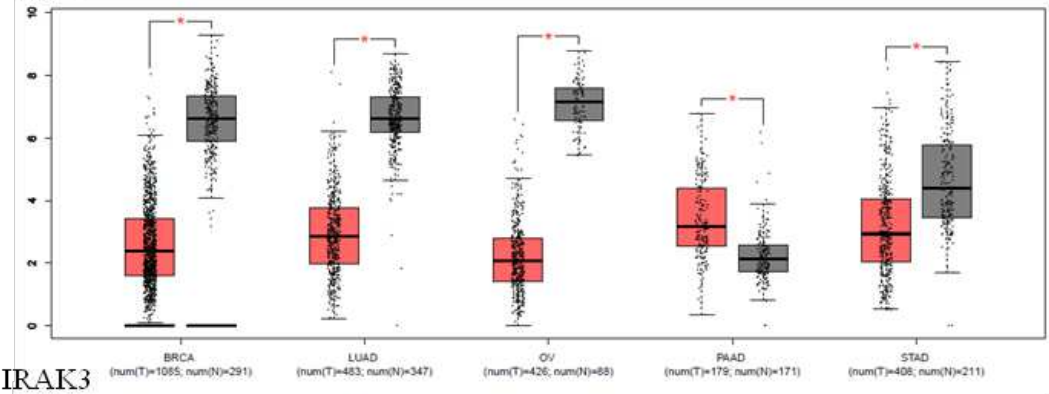
TISIDB


Ovarian
(2222)
1195
Breast
(2349)
Pancreatic
(2349)
1263
1206
175
175
172
24
29
33
17
13
171
59
43
273
3
12
193
58
21
46
172
38
14
28
146
203
51
190
1340
1344
Gastric
(2390)
Lung
(2328)

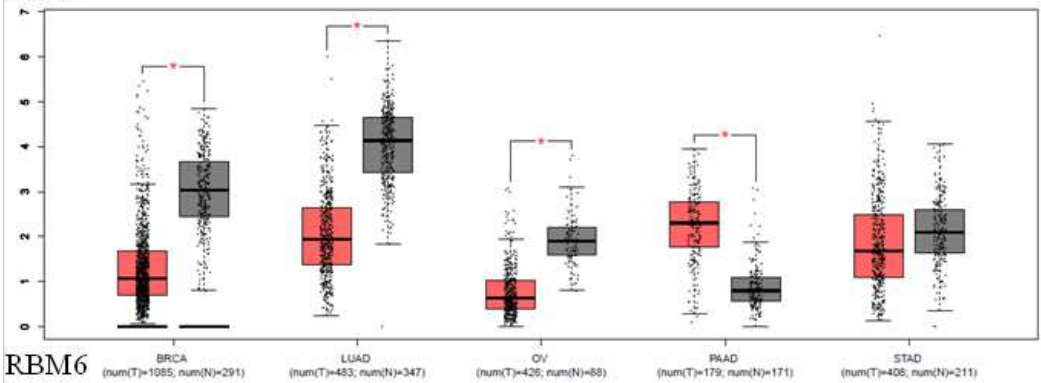

Molecular Function
NAD-retinol Dehydrogenase Activity (GO:0004745)
RNA Binding (GO:0003723)
Transcription Regulatory Region Nucleic Acid Binding (GO:0001067)
Protease Binding (GO:0002020)
Serine-Type Endopeptidase Inhibitor Activity (GO:0004867)
DNA Binding (GO:0003677)
4-galactosyl-N-acetylglucosaminide 3-alpha-L-fucosyltransferase Activity (GO:0017083)
ADP-D-ribose Modification-Dependent Protein Binding (GO:0160002)
Calcium-Dependent Protein Serine/Threonine Phosphatase Activity (GO:0004723)
Nucleosome Binding (GO:0031491)
Biological process
Positive Regulation Of Biosynthetic Process (GO:0009891)
Positive Regulation Of Myelination (GO:0031643)
Myelination In Peripheral Nervous System (GO:0022011)
Negative Regulation Of Transmembrane Receptor Protein Serine/Threonine Kinase Signaling Pathway (GO:0090101)
Regulation Of Transforming Growth Factor Beta Receptor Signaling Pathway (GO:0017015)
Positive Regulation Of MAPK Cascade (GO:0043410)
Myeloid Cell Differentiation (GO:0030099)
Regulation Of Apoptotic Process (GO:0042981)
Positive Regulation Of Extracellular Matrix Organization (GO:1903055)
Positive Regulation Of Nervous System Process (GO:0031646)
KEGG pathway
Axon guidance
Rheumatoid arthritis
NF-kappa B signaling pathway
Rap1 signaling pathway
TNF signaling pathway
Autophagy
Notch signaling pathway
MAPK signaling pathway
Retinol metabolism
Adipocytokine signaling pathway

Gene	*Lung*	*Ovarian*	*Breast*	*Pancreatic*	*Gastric*
TNXB	–	↘	↗	↗	↗
IRAK3	↗	↘	↘	–	↘
RBM6	↗	–	↘	↗	–
SERPINA1	–	↘	↗	↗	↗
SUB1	↗	↘	–	↗	↘

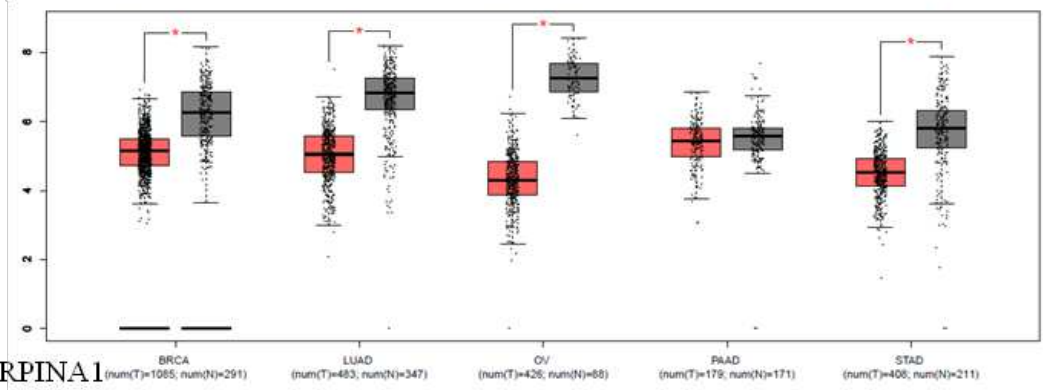
TNXB



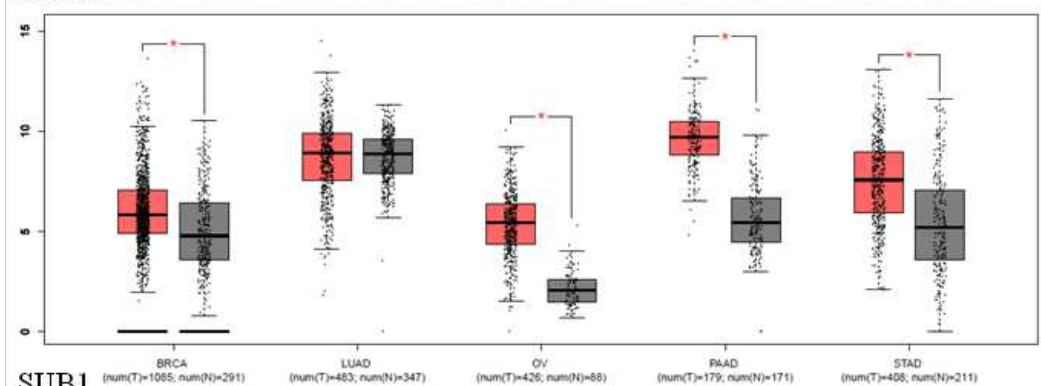
IRAK3



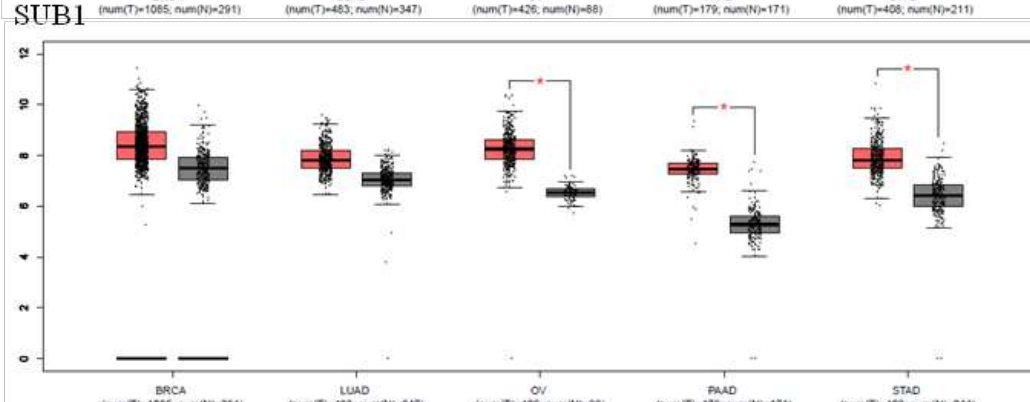
RBM6

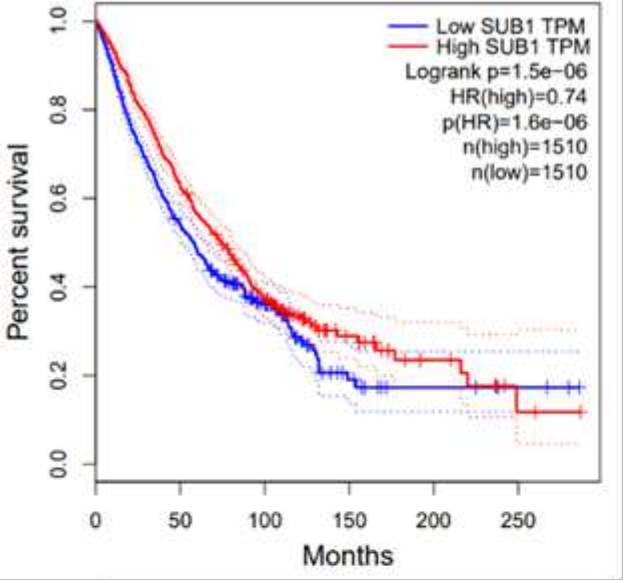


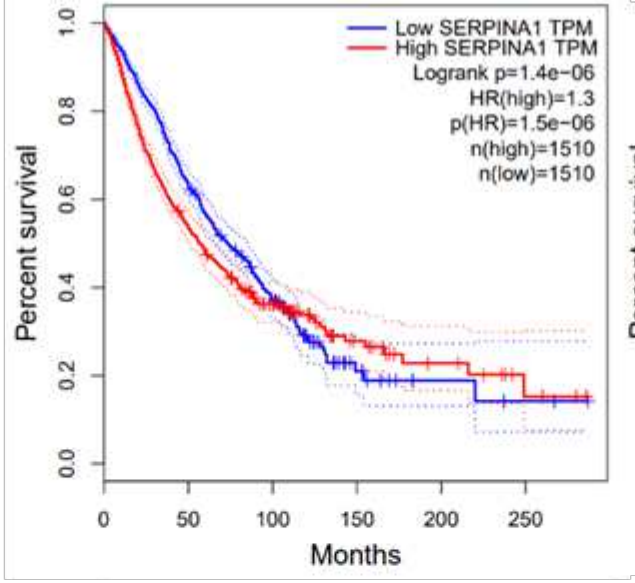
SERPINA1

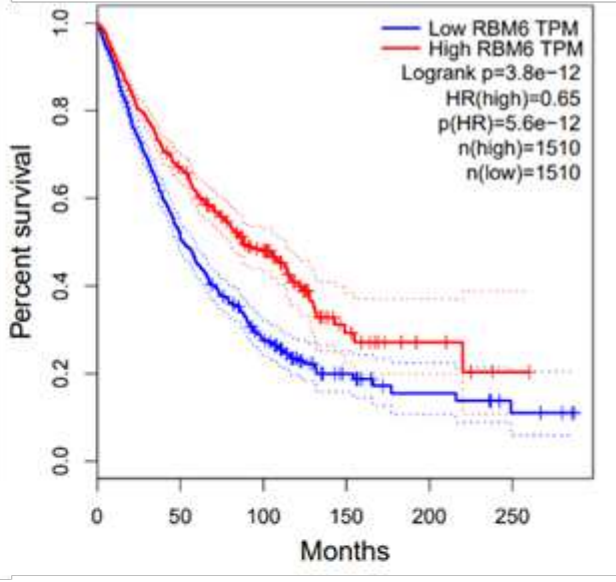


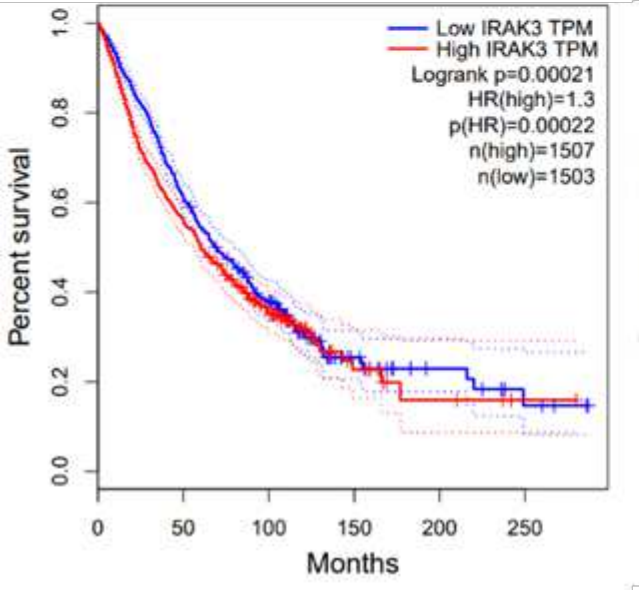
SUB1

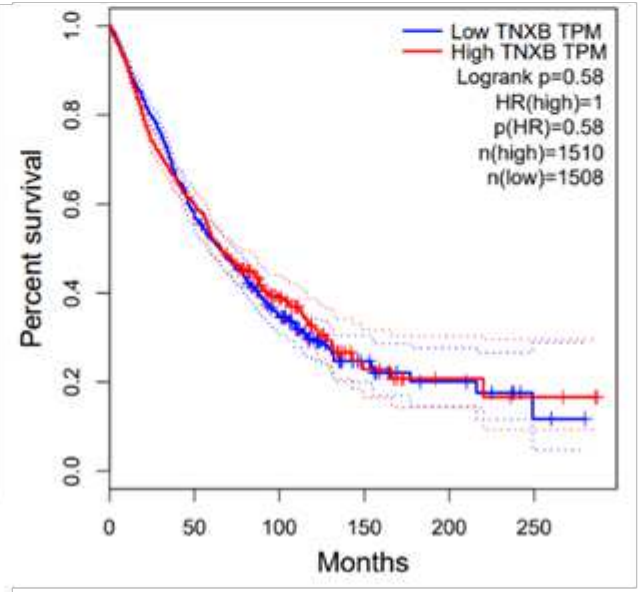


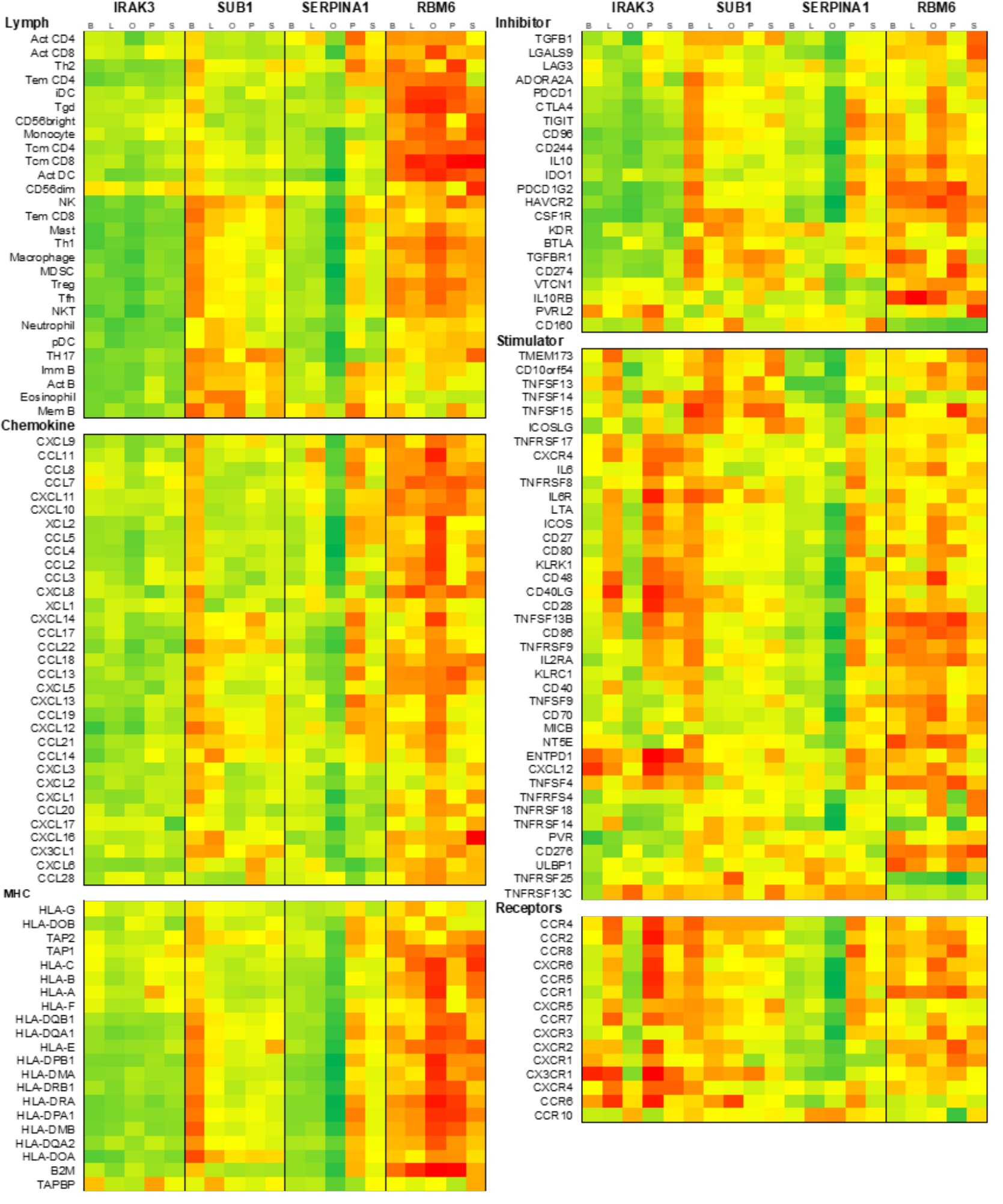

Low SUB1 TPM
High SUB1 TPM
Logrank p=1.5e−06
HR(high)=0.74
p(HR)=1.6e−06
n(high)=1510
n(low)=1510
Percent survival
Months

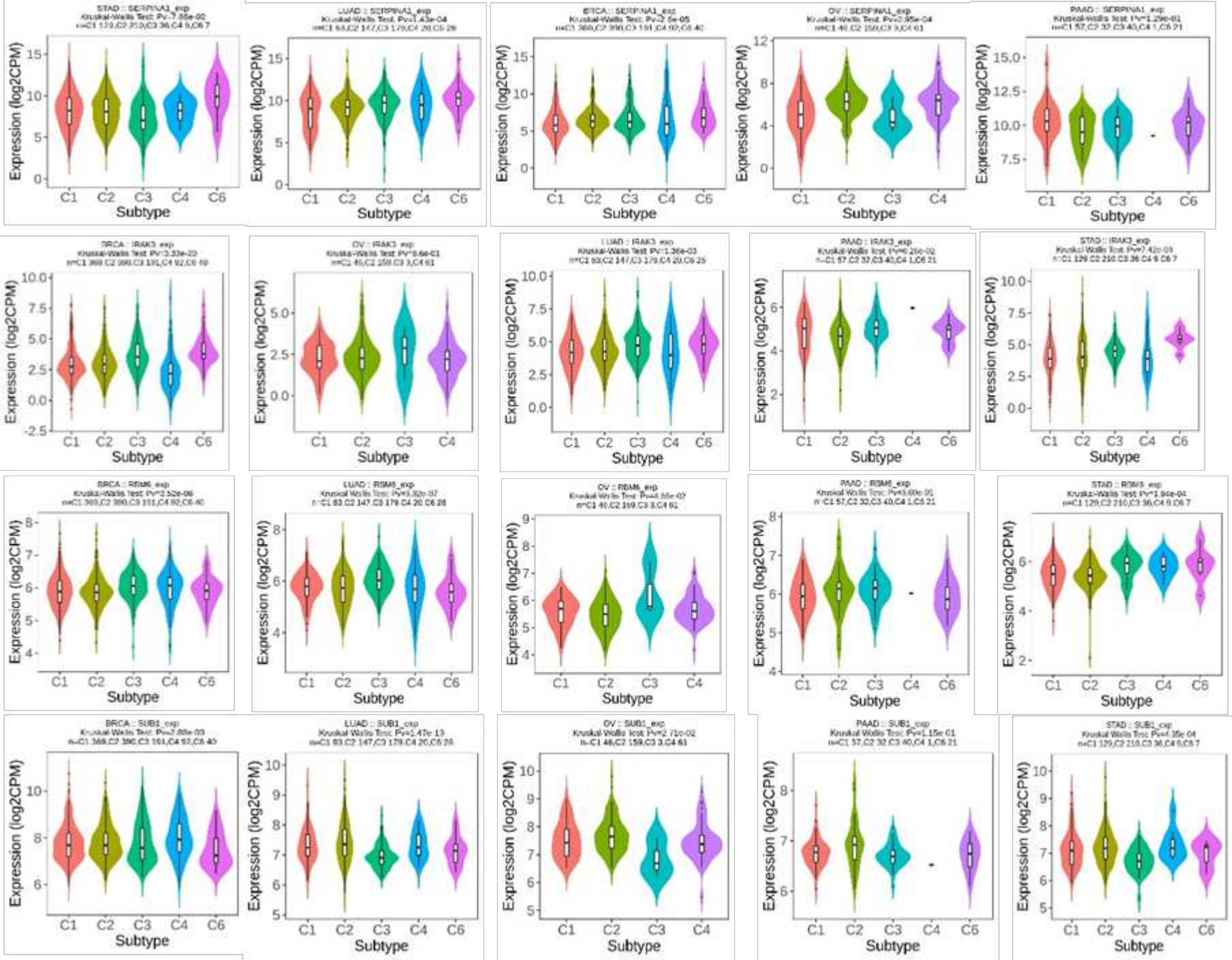

Low SERPINA1 TPM
High SERPINA1 TPM
Logrank p=1.4e−06
HR(high)=1.3
p(HR)=1.5e−06
n(high)=1510
n(low)=1510
Percent survival
Months


Low RBM6 TPM
High RBM6 TPM
Logrank p=3.8e−12
HR(high)=0.65
p(HR)=5.6e−12
n(high)=1510
n(low)=1510
Percent survival
Months


Low IRAK3 TPM
High IRAK3 TPM
Logrank p=0.00021
HR(high)=1.3
p(HR)=0.00022
n(high)=1507
n(low)=1503
Percent survival
Months


Low TNXB TPM
High TNXB TPM
Logrank p=0.58
HR(high)=1
p(HR)=0.58
n(high)=1510
n(low)=1508
Percent survival
Months


IRAK3
SUB1
SERPINA1
RBM6
B L O P S
Lymph
Act CD4
Act CD8
Th2
Tem CD4
iDC
Tgd
CD56bright
Monocyte
Tcm CD4
Tcm CD8
Act DC
CD56dim
NK
Tem CD8
Mast
Th1
Macrophage
MDSC
Treg
Tfh
NKT
Neutrophil
pDC
TH17
Imm B
Act B
Eosinophil
Mem B
Chemokine
CXCL9
CCL11
CCL8
CCL7
CXCL11
CXCL10
XCL2
CCL5
CCL4
CCL2
CCL3
CXCL8
XCL1
CXCL14
CCL17
CCL22
CCL18
CCL13
CXCL5
CXCL13
CCL19
CXCL12
CCL21
CCL14
CXCL3
CXCL2
CXCL1
CCL20
CXCL17
CXCL16
CX3CL1
CXCL6
CCL28
MHC
HLA-G
HLA-DOB
TAP2
TAP1
HLA-C
HLA-B
HLA-A
HLA-F
HLA-DQB1
HLA-DQA1
HLA-E
HLA-DPB1
HLA-DMA
HLA-DRB1
HLA-DRA
HLA-DPA1
HLA-DMB
HLA-DQA2
HLA-DOA
B2M
TAPBP
Inhibitor
TGFB1
LGALS9
LAG3
ADORA2A
PDCD1
CTLA4
TIGIT
CD96
CD244
IL10
IDO1
PDCD1G2
HAVCR2
CSF1R
KDR
BTLA
TGFBR1
CD274
VTCN1
IL10RB
PVRL2
CD160
Stimulator
TMEM173
CD10orf54
TNFSF13
TNFSF14
TNFSF15
ICOSLG
TNFRSF17
CXCR4
IL6
TNFRSF8
IL6R
LTA
ICOS
CD27
CD80
KLRK1
CD48
CD40LG
CD28
TNFSF13B
CD86
TNFRSF9
IL2RA
KLRC1
CD40
TNFSF9
CD70
MICB
NT5E
ENTPD1
CXCL12
TNFSF4
TNFRFS4
TNFRSF18
TNFRSF14
PVR
CD276
ULBP1
TNFRSF25
TNFRSF13C
Receptors
CCR4
CCR2
CCR8
CXCR6
CCR5
CCR1
CXCR5
CCR7
CXCR3
CXCR2
CXCR1
CX3CR1
CXCR4
CCR6
CCR10


Expression (log2CPM)
Subtype
C1
C2
C3
C4
C6

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryfile1.xlsx
- supplementaryfile2.xlsx