Original Article

sciendo

Pan-cancer landscape of the RUNX protein family reveals their potential as carcinogenic biomarkers and the mechanisms underlying their action

Shen Pan[1], Siyu Sun[2], Bitian Liu[3], Yang Hou[1]

[1]Department of Radiology, Shengjing Hospital of China Medical University, Shenyang 110004, Liaoning Province, China; [2]Department of Gastroenterology, Shengjing Hospital of China Medical University, Shenyang 110004, Liaoning Province, China; [3]Department of Urology, Shengjing Hospital of China Medical University, Shenyang 110004, Liaoning Province, China

ABSTRACT

Background: The RUNX family of transcription factors plays an important regulatory role in tumor development. Although the importance of RUNX in certain cancer types is well known, the pan-cancer landscape remains unclear. **Materials and Methods**: Data from The Cancer Genome Atlas (TCGA) provides a pan-cancer overview of the *RUNX* genes. Hence, herein, we performed a pan-cancer analysis of abnormal *RUNX* expression and deciphered the potential regulatory mechanism. Specifically, we used TCGA multi-omics data combined with multiple online tools to analyze transcripts, genetic alterations, DNA methylation, clinical prognoses, miRNA networks, and potential target genes. **Results**: *RUNX* genes are consistently overexpressed in esophageal, gastric, pancreatic, and pan-renal cancers. The total protein expression of *RUNX1* in lung adenocarcinoma, kidney renal clear cell carcinoma (KIRC), and uterine corpus endometrial carcinoma (UCEC) is consistent with the mRNA expression results. Moreover, increased phosphorylation on the T14 and T18 residues of RUNX1 may represent potential pathogenic factors. The *RUNX* genes are significantly associated with survival in pan-renal cancer, brain lower-grade glioma, and uveal melanoma. Meanwhile, various mutations and posttranscriptional changes, including the *RUNX1* D96 mutation in invasive breast carcinoma, the co-occurrence of *RUNX* gene mutations in UCEC, and methylation changes in the *RUNX2* promoter in KIRC, may be associated with cancer development. Finally, analysis of epigenetic regulator co-expression, miRNA networks, and target genes revealed the carcinogenicity, abnormal expression, and direct regulation of *RUNX* genes. **Conclusions**: We successfully analyzed the pan-cancer abnormal expression and prognostic value of *RUNX* genes, thereby providing potential biomarkers for various cancers. Further, mutations revealed via genetic alteration analysis may serve as a basis for personalized patient therapies.

Key words: *RUNX* family, The Cancer Genome Atlas, pan-cancer analysis, prognosis, regulatory mechanism

Address for Correspondence:
Yang Hou, Department of Radiology, Shengjing Hospital of China Medical University, No. 36 Sanhao Street, Heping District, Shenyang 110004, Liaoning Province, China.
E-mail: houyang1973@163.com

Bitian Liu, Department of Urology, Shengjing Hospital of China Medical University, No. 36 Sanhao Street, Heping District, Shenyang 110004, Liaoning Province, China.
E-mail: btliu@cmu.edu.cn

Access this article online

Website: www.intern-med.com

DOI: 10.2478/jtim-2022-0013

INTRODUCTION

Transcription factors (TFs), which account for approximately 8% of all human genes, function as master regulators in multiple signaling pathways in eukaryotic cells. In fact, approximately 20% of TFs are associated with at least one human phenotype.[1] However, TF regulation is highly dynamic, where only a single TF can control the transcription of multiple genes in different cell types. Therefore, dysregulated TFs contribute to the pathogenesis of many maladies, including cardiovascular diseases, inflammatory diseases, and various cancers,[2] thus highlighting the physiological importance of the gene regulatory mechanisms mediated by TFs. The activity of TFs can be directly modulated by numerous mechanisms, including point

mutations, gene amplification or deletion, and disordered expression. Moreover, indirect changes in noncoding DNA mutations may affect the activity of TFs.[3] Therefore, considering that TFs are gradually being considered as cancer treatment targets,[3,4] characterizing TF abnormalities within a pan-cancer context using a multi-omics approach, although daunting, is necessary. Particularly, these findings would provide insights regarding TF pathophysiology, while helping to decode specific functions of the genome.

The *RUNX* family comprises metazoan TFs, which serve as the primary development regulators. However, they are often dysregulated in human cancers and play an important, and often contradictory, role in cancer pathogenesis.[5] Three RUNX proteins, namely, RUNX1, RUNX2, and RUNX3, exhibit different tissue-specific expression profiles in mammals[6] and possess the Runt and Runx inhibition (RunxI) domains. The Runt domain is an evolutionarily conserved protein domain responsible for DNA-binding and protein–protein interactions.[7,8] Moreover, RUNX-related TFs contain a RUNXI domain at the C-terminus, which might interact with functional cofactors.[9] Meanwhile, RUNX1-related somatic mutations and chromosomal rearrangements are frequently observed in myeloid and lymphoid leukemia cells.[10] Similarly, aberrant expression of RUNX2 is a key pathological feature of osteosarcoma.[11] However, to the best of our knowledge, pan-cancer research on RUNX2 has not yet been reported.

In this study, we used data from The Cancer Genome Atlas (TCGA) to perform a comprehensive pan-cancer analysis of *RUNX* genes. Our primary aim was to decipher the commonalities between the various *RUNX* genes in cancer, while deciphering the mechanisms underlying their aberrant expression. To this end, we analyzed *RUNX* genes based on the transcript data, protein expression, genetic alteration, DNA methylation, clinical prognosis, miRNA networks, and related functions. We detected that abnormal *RUNX* expression in certain cancers impacts disease prognosis. Moreover, genetic alterations, changes in DNA methylation levels, and co-expression with *RUNX* genes may represent potential pathogenic factors. Construction of a miRNA network and identification of putative target genes enabled the generation of a regulatory landscape for the abnormal expression of *RUNX* genes.

MATERIALS AND METHODS

mRNA expression analysis

RNAseq-HTSeq-FPKM data from 33 types of cancer were downloaded from UCSC Xena (xenabrowser.net). We then compared the expression of *RUNX* family genes between tumor and healthy tissues using TCGA data. ggpubr R package in R 3.6.2 (www.r-project.org) was used to generate a pan-cancer Box plot for RUNX expression. As certain tumor samples contain very limited healthy tissues, we obtained corresponding normal tissues from the genotype-tissue expression (GTEx) database for expression analysis on GEPIA2 (http://gepia2.cancer-pku.cn/#analysis), under the default parameters of *P*-value <0.01, $|\log_2 FC| > 1$, and "Match TCGA normal and GTEx data." Finally, we set the number of normal tissues to more than five as a valid comparison. To compare the tissue expression specificity and gene similarity in the *RUNX* gene family, we obtained the gene expression cluster heatmap of various tissues from GTEx (www.gtexportal.org).

Protein expression analysis

We explored the abundance of *RUNX* family proteins through the Clinical Proteomic Tumor Analysis Consortium (CPTAC) database in the UALCAN portal (ualcan.path.uab.edu/analysis-prot.html).[12] Specifically, we performed differential analysis between the normal and tumor tissues for six tumor types: breast cancer, ovarian cancer, colon cancer, kidney renal clear cell carcinoma (KIRC), uterine corpus endometrial carcinoma (UCEC), and lung adenocarcinoma (LUAD). In addition to total protein content, phosphoproteins at different sites were also analyzed. The T14, S21, T18, S212, T14S17, and T14S21 residues are phosphorylated in RUNX1 (NP_001001890.1). The S28, S275, and S340 residues are phosphorylated in RUNX2 (NP_001019801.3). The T28, S228, and T28S31 residues are phosphorylated in RUNX3 (NP_001026850.1).

Survival prognosis analysis

GEPIA2 was employed to determine overall survival (OS) and disease-free survival (DFS) using TCGA tumor data.[13] The median group cutoff was considered the threshold for grouping high and low expression. Survival plots were obtained through the "Survival Analysis" module, using the log-rank test.

Genetic alteration analysis

The mutation type and copy number alteration in all tumors (TCGA) were summarized using cBioPortal (www.cbioportal.org).[14] After selecting TCGA Pan Cancer Atlas Studies based on the *RUNX* family, the genetic alteration characteristics, sorted by alteration frequency, were displayed in the "Cancer Types Summary" module. We used the "Mutations" module to depict the mutation sites in the *RUNX* family proteins (illustrated as a schematic of the protein structure). After identifying cases that carried mutations, the distribution of cases with mutation co-occurrence was assessed. Tumor mutation burden (TMB) based on VarScan2 variant data of somatic mutations from UCSC Xena was calculated as the total mutation

incidences per million base pairs. After calculating the correlation between the expression of *RUNX* genes and TMB, we used the fmsb R package to construct a radar chart illustrating the pan-cancer relationships between *RUNX* genes and TMB.

Fusion gene analysis

We downloaded the fusion gene data of the *RUNX* family from TCGA Fusion Gene Database (http://www.tumorfusions.org/), including that of the fusion genes associated with the 33 cancer types (predicted by PRADA).

Methylation analysis

Data related to differential methylation, methylation survival, and methylation to the expression of RUNX family genes were obtained from GSCALite (bioinfo.life.hust.edu.cn/web/GSCALite/).[15] Differential methylation analysis was based on 14 cancer types and included paired tumor versus normal data. To identify the genes that were significantly influenced by genome methylation, we analyzed the association between paired mRNA expression and methylation based on Pearson's product-moment correlation coefficients. Methylation and clinical OS data were combined, and median methylation level was used to divide gene methylation into two groups. Cox regression was then performed to estimate the hazards (risk of death). The high-methylation group exhibited a worse survival when the Cox coefficient was > 0. Thus, hyper-worse survival was defined as "High," and otherwise it was defined as "Low."

To accurately explore the effect of methylation sites, particularly those upstream of the transcription start site (TSS), on *RUNX2* expression and its relationship with survival prognosis, we used MEXPRESS and MethSurv online analysis tools. MEXPRESS (mexpress.be) is an online tool that integrates gene expression, DNA methylation, and clinical data in TCGA.[16] It allows us to visualize the relationship between specific methylation sites and gene expression levels. Meanwhile, MethSurv (biit.cs.ut.ee/methsurv/) is an online tool for performing survival analysis based on specific methylation sites by employing methylome data from TCGA.[17]

Epigenetic regulators

To explore the relationship between *RUNX* and epigenetics in various cancers, we collected genes related to epigenetic processes, such as chromatin remodeling, DNA methylation, histone acetylation, histone methylation, readers, and epigenetic mediators.[18–20] Using the reshape2 and RColorBrewer R packages, we calculated the co-expression of the *RUNX* genes and epigenetic regulators based on TCGA pan-cancer RNAseq-HTSeq-FPKM data and visualized the results on a heatmap.

Pathway analysis

The "Pathway Activity" module of GSCALite was used to investigate the correlation between *RUNX* family members and 10 well-known cancer-related pathways, namely, TSC/mTOR, RTK, RAS/MAPK, PI3K/AKT, hormone ER, hormone AR, epithelial–mesenchymal transition (EMT), DNA damage response, cell cycle, and apoptosis pathways.[15] Pathway scores were calculated based on TCGA pan-cancer data.[21] Global percentage presents the percentage of 32 cancer types in which a gene affects the cancer pathway. Alternatively, heatmaps show proteins that are inhibited or activated in at least five cancer types.

Stromal and immune infiltration analysis

Stromal and immune fractions were evaluated using the Estimation of STromal and Immune cells in MAlignant Tumor tissues using expression data (ESTIMATE) program.[22] Using the estimate.R package in R 3.6.2, tumor stromal and immune infiltration of TCGA pan-cancer samples was calculated using the profiles of two gene sets that included 141 genes.

miRNA network

The miRNA regulation analysis was conducted using GSCALite, which is based on verified experimental (TarBase, miRTarBase, and mir2disease) and predicted databases (targetscan and miRanda).[15] A correlation analysis of *RUNX* gene expression between paired mRNA and miRNA was performed to explore the miRNA–gene regulatory network in all the 33 cancer types. A Pearson correlation with an $FDR \leq 0.05$ and an $R < 0$ was considered a significant negative correlation.

Transcriptional regulation of target genes

We used Harmonizome (amp.pharm.mssm.edu/Harmonizome/) to explore the pan-cancer targets of the *RUNX* genes. Harmonizome is a collection of processed datasets containing gene and protein interaction data.[23] After entering *RUNX* genes, we selected the following four seed databases: CHEA, JASPAR, TRANSFAC Curated, and TRANSFAC, and subsequently predicted potential *RUNX* target genes. Using the reshape2 and RColorBrewer R packages, we identified pan-cancer correlations between *RUNX* genes and potential target genes, which was based on TCGA pan-cancer RNAseq-HTSeq-FPKM data. The results were visualized on a heatmap.

RESULTS

Expression of RUNX genes in various tumors

The family of *RUNX* genes is abnormally expressed in most cancers. In particular, *RUNX1* exhibited significant differential overexpression in all assessed cancers (except in prostate adenocarcinoma [PRAD]). *RUNX2* and *RUNX1* have similar cancer expression characteristics (Figure 1A, B);

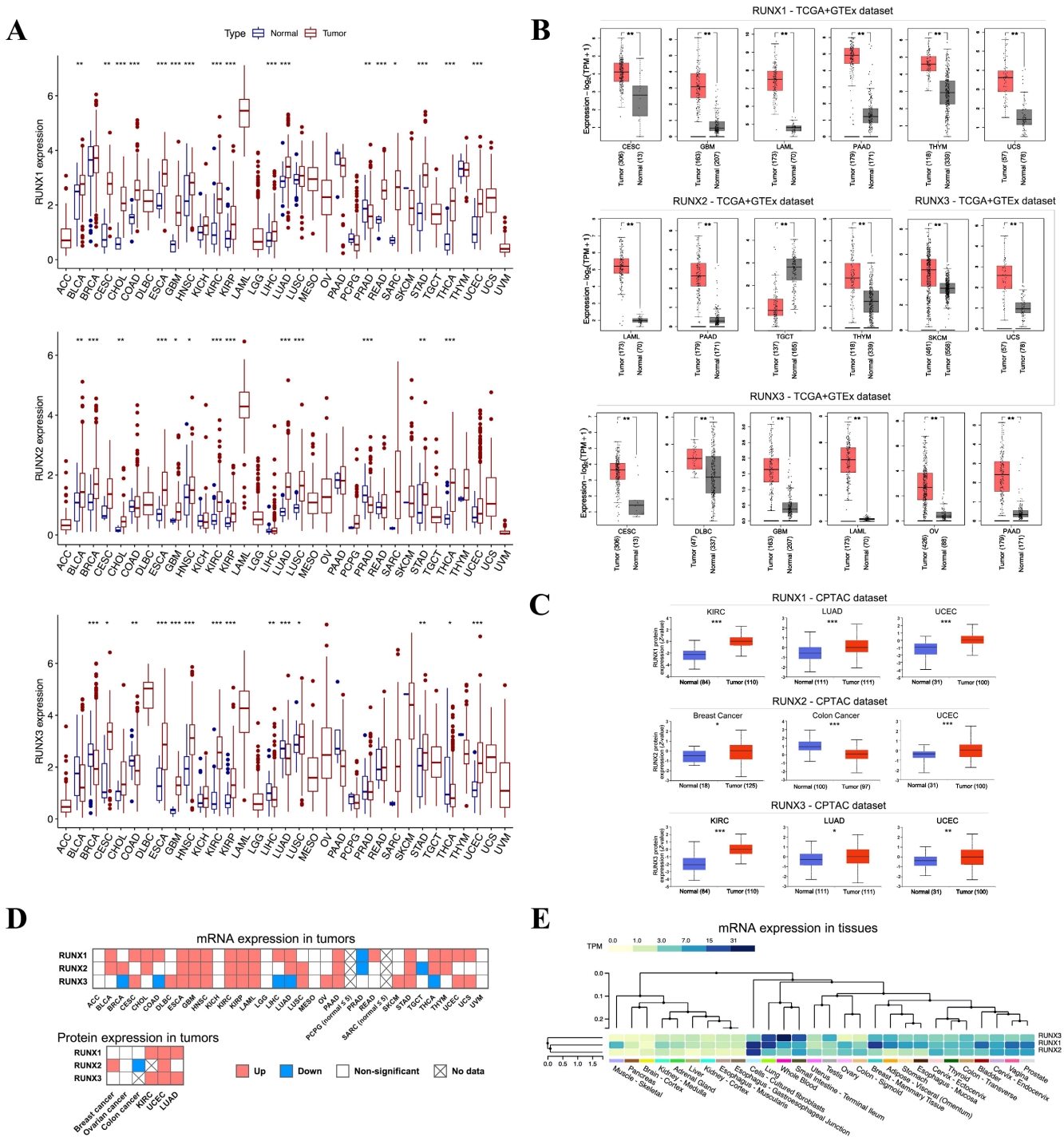

Figure 1: Pan-cancer expression of *RUNX* genes. (A) Comparison of *RUNX* mRNA expression in cancer and adjacent normal tissues. (B) Comparison of *RUNX* expression data from TCGA and GTEx databases in cancers that lacked or had fewer normal tissues. (C) Total RUNX protein expression in certain cancers. (D) The pan-cancer landscape of *RUNX* expression at mRNA and protein levels. (E) Heatmap of *RUNX* gene expression in various tissues. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. TCGA: The Cancer Genome Atlas; GTEx: genotype-tissue expression; BLCA: bladder urothelial carcinoma; BRCA: breast carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD: colon adenocarcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: esophageal carcinoma; GBM: glioblastoma multiforme; HNSC: head and neck squamous cell carcinoma; KICH: kidney chromophobe; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LAML: acute myeloid leukemia; LGG: lower-grade glioma; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; PAAD: pancreatic adenocarcinoma; PRAD: prostate adenocarcinoma; SKCM: skin cutaneous melanoma; STAD: stomach adenocarcinoma; TGCT: testicular germ cell tumors; THCA: thyroid carcinoma; THYM: thymoma; UCEC: uterine corpus endometrial carcinoma; UVM: uveal melanoma.

however, the former exhibited low expression in testicular germ cell tumors (TGCT). Meanwhile, the pan-cancer expression trend of *RUNX3* differed from those of *RUNX1* and *RUNX2,* with *RUNX3* found downregulated in invasive breast carcinoma (BRCA), colon adenocarcinoma (COAD), liver hepatocellular carcinoma (LIHC), LUAD, and thymoma (THYM; Figure 1A, B).

Collectively, *RUNX* genes were overexpressed in eight cancers, namely, esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), KIRC, kidney renal papillary cell carcinoma (KIRP), acute myeloid leukemia (LAML), pancreatic adenocarcinoma (PAAD), and stomach adenocarcinoma (STAD). Among the six adult tumors in the CPTAC database, *RUNX1*, *RUNX2*, and *RUNX3* independently influenced three tumor types via significant protein expression (Figure 1C). These TCGA pan-cancer analysis results were summarized using Prism 8 version 8.4.0. The corresponding mRNA and protein expression levels are depicted in heatmaps (Figure 1D).

RUNX gene expression in different tissues

The expression level of *RUNX* family genes varied between tissues. Based on cluster analysis, the expression of *RUNX1* and *RUNX2* was most similar. Moreover, *RUNX* family genes exhibited relatively low expression in the brain, esophagus, liver, kidney, pancreas, and adrenal gland and relatively high expression in the fibroblasts, lungs, and blood (Figure 1E).

RUNX protein phosphorylation

In the CPTAC database, the phosphorylation levels at different RUNX residues showed consistent trends for all phosphorylation sites (Figure 2A). All phosphorylation sites are located outside the Runt and RunxI domains (Figure 2B). To analyze CPTAC-identified phosphorylation of RUNX family proteins, the PhosphoNET database (www.phosphonet.ca) was used to analyze the evolutionary conservation of these phosphorylation sites. The most functionally important phosphosites are expected to be highly conserved. Importantly, with the exception of the T28 residue of RUNX3, all other phosphorylated residues have been identified in mammals. However, the highly conserved T14 and T18 residues of RUNX1 might have an important role in various cancers. Meanwhile, upon considering the influence of the neighboring phosphosites, the Kinase Predictor module identified extracellular regulated protein kinases 1 (ERK1) and ERK2 as the proteins that are most likely to phosphorylate all of the residues described above.

Survival analysis

In GEPIA2, we analyzed the effect of RUNX family members on the OS and DFS of 33 cancers. With the exception of BRCA, upregulation of *RUNX1* was associated with a poor prognosis and was significantly associated with poor OS and DFS (Figure 3A). Similarly, *RUNX2* exhibited consistently high pan-cancer expression and was associated with poor prognosis (Figure 3B). Meanwhile, *RUNX3* expression was noticeably different from that of *RUNX1* and *RUNX2* in that it was associated with significant survival outcomes only in individual cancers (Figure 3C). The pan-cancer survival heatmap revealed that brain lower-grade glioma (LGG) exhibited significant, and consistent, survival outcomes in RUNX family analysis (Figure 3D). High *RUNX1* and *RUNX2* was associated with poor uveal melanoma (UVM) prognosis. Moreover, pan-renal cancer analysis suggested that high *RUNX1* and *RUNX2* expression was associated with poor prognosis (Figure 3D).

Genetic alteration summary analysis

RUNX1 had the highest alteration frequency in LAML, with an approximate 10% mutation frequency and 4% gene fusion. Further, *RUNX1* had the second highest alteration frequency in ESCA, with >6% deep deletion (Figure 4A). Meanwhile, *RUNX2* had the highest alteration frequency in ESCA, with up to 6% amplification. Indeed, among the 33 cancer types, genetic alterations in *RUNX2* were primarily associated with amplification (Figure 4B). However, although the overall genetic alteration frequency of *RUNX3* was lower than those of *RUNX1* and *RUNX2*, it had the highest alteration frequency (<4%) in skin cutaneous melanoma (SKCM; Figure 4C).

Using the data in cBioPortal, we further analyzed cancers with a relatively high frequency of single-type genetic alteration. We found that missense mutations caused an increase in *RUNX1* expression in LAML samples (Figure 4D). Although deep deletion in ESCA caused a decrease in *RUNX2* expression (Figure 4D), there was no significant change in *RUNX1* expression (Figure 1D). Meanwhile, in UCEC, genetic alterations in the *RUNX* family ranked among the top five, of which *RUNX1* and *RUNX2* accounted for approximately 5%; however, this did not affect the RUNX family gene expression (Figure 4D, E). ESCA and STAD, which are similar in terms of anatomical structure, also exhibited *RUNX2* amplification; however, this did not contribute to an increase in *RUNX2* expression (Figure 4E).

Mutation analysis

Using the "Mutations" module in the cBioPortal, we depicted all mutation sites in the *RUNX* family and annotated the hotspots. More truncating mutations were identified in *RUNX1* than in *RUNX2* and *RUNX3,* while the D96 site marked by the hotspot in the Runt domain

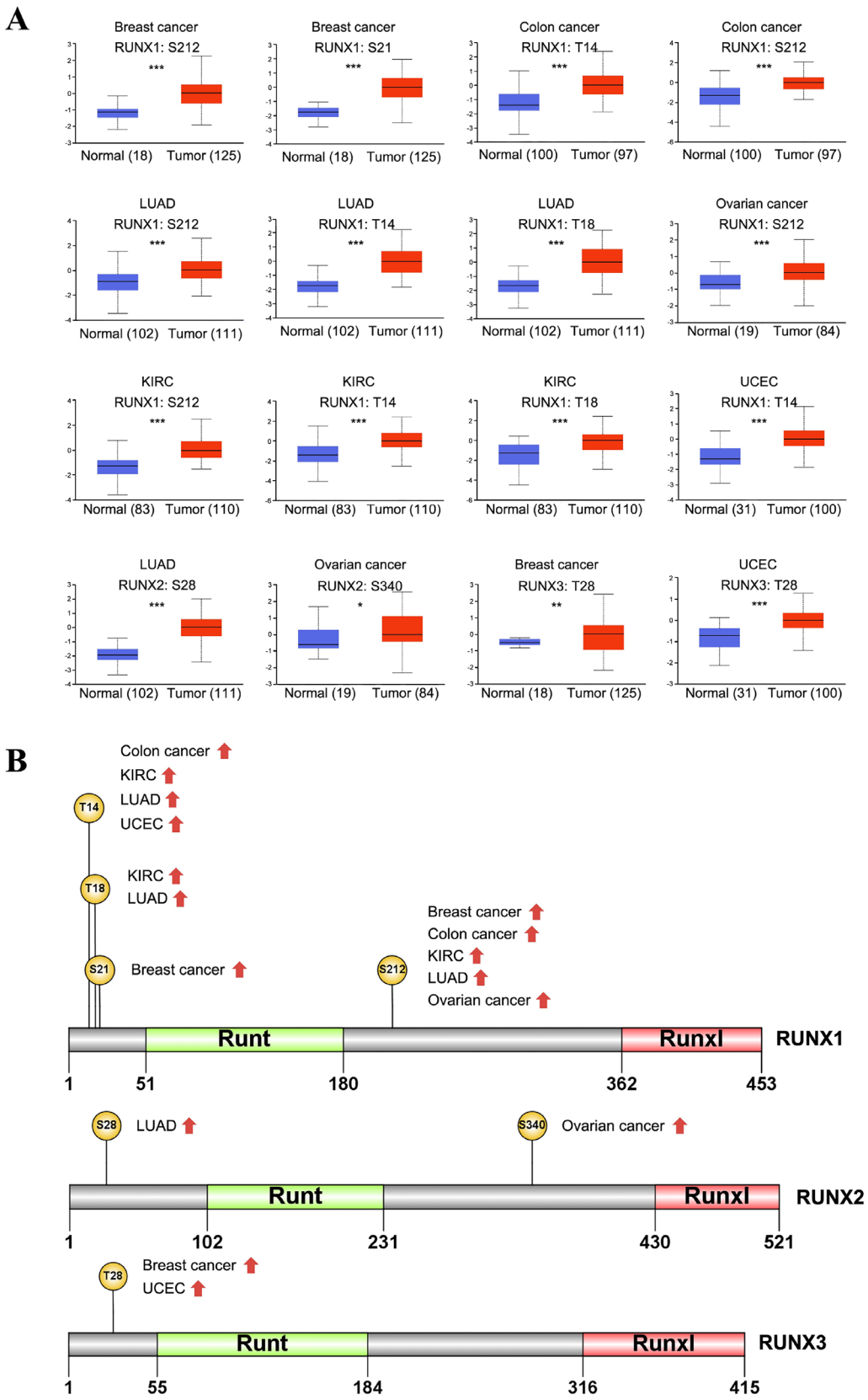

Figure 2: Phosphorylation analysis of RUNX proteins in different tumors. (A) Box plots show the phosphorylation sites in RUNX proteins that are positively associated with different cancers. (B) The RUNX protein schematic shows phosphoprotein sites with positive results. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. KIRC: kidney renal clear cell carcinoma; LUAD: lung adenocarcinoma; UCEC: uterine corpus endometrial carcinoma.

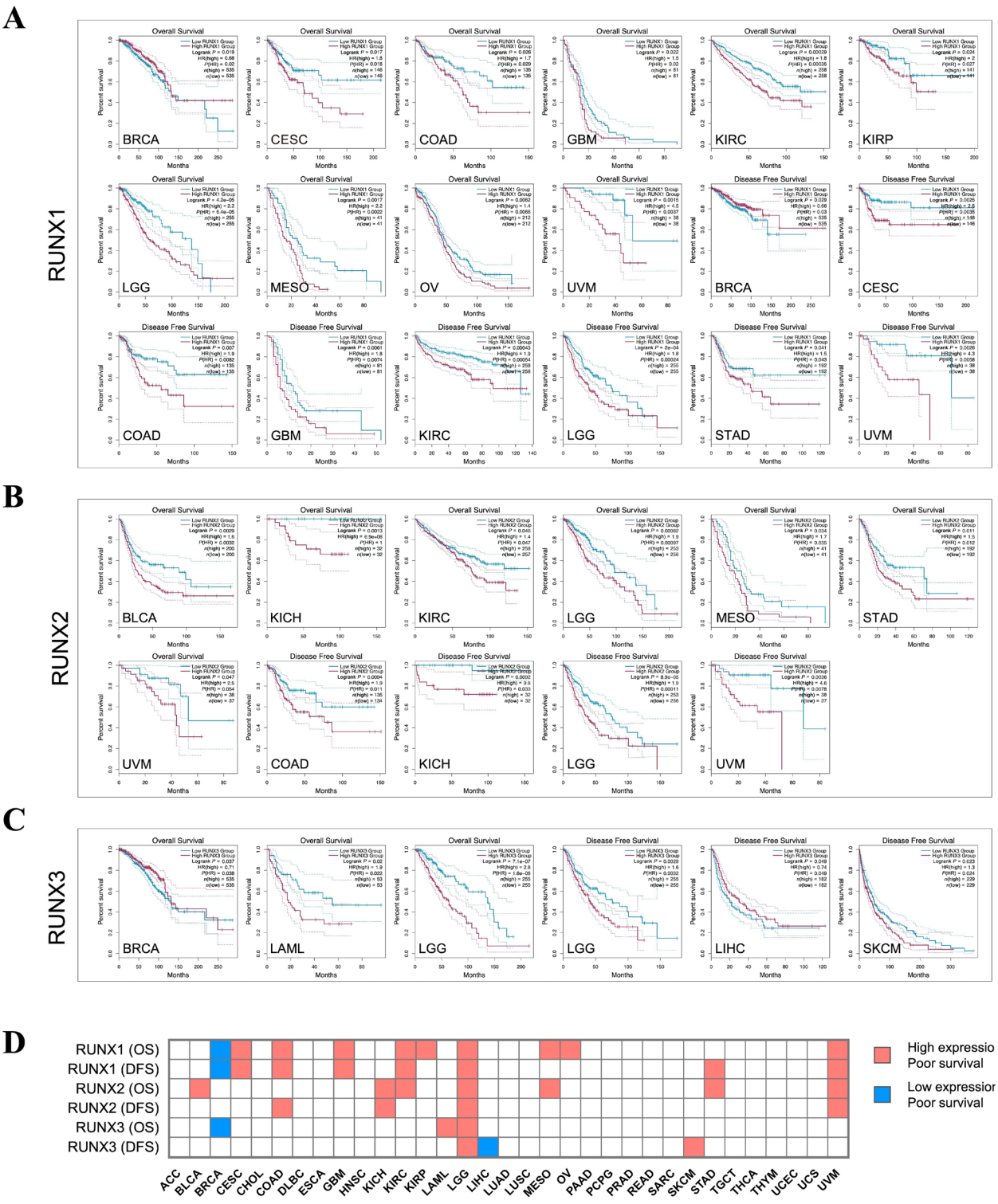

Figure 3: *RUNX*-associated survival analyses. Positive association of (A) *RUNX1*, (B) *RUNX2*, and (C) *RUNX3* with OS and DFS. (D) The pan-cancer survival landscape of *RUNX* genes. OS: overall survival; DFS: disease-free survival; BRCA: breast carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD: colon adenocarcinoma; GBM: glioblastoma multiforme; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LGG: lower-grade glioma; UVM: uveal melanoma; BRCA: breast carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD: colon adenocarcinoma; GBM: glioblastoma multiforme; KIRC: kidney renal clear cell carcinoma; STAD: stomach adenocarcinoma; BLCA: bladder urothelial carcinoma; KICH: kidney chromophobe; SKCM: skin cutaneous melanoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: esophageal carcinoma; HNSC: head and neck squamous cell carcinoma; LAML: acute myeloid leukemia; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; PAAD: pancreatic adenocarcinoma; PRAD: prostate adenocarcinoma; TGCT: testicular germ cell tumors; THCA: thyroid carcinoma; THYM: thymoma; UCEC: uterine corpus endometrial carcinoma.

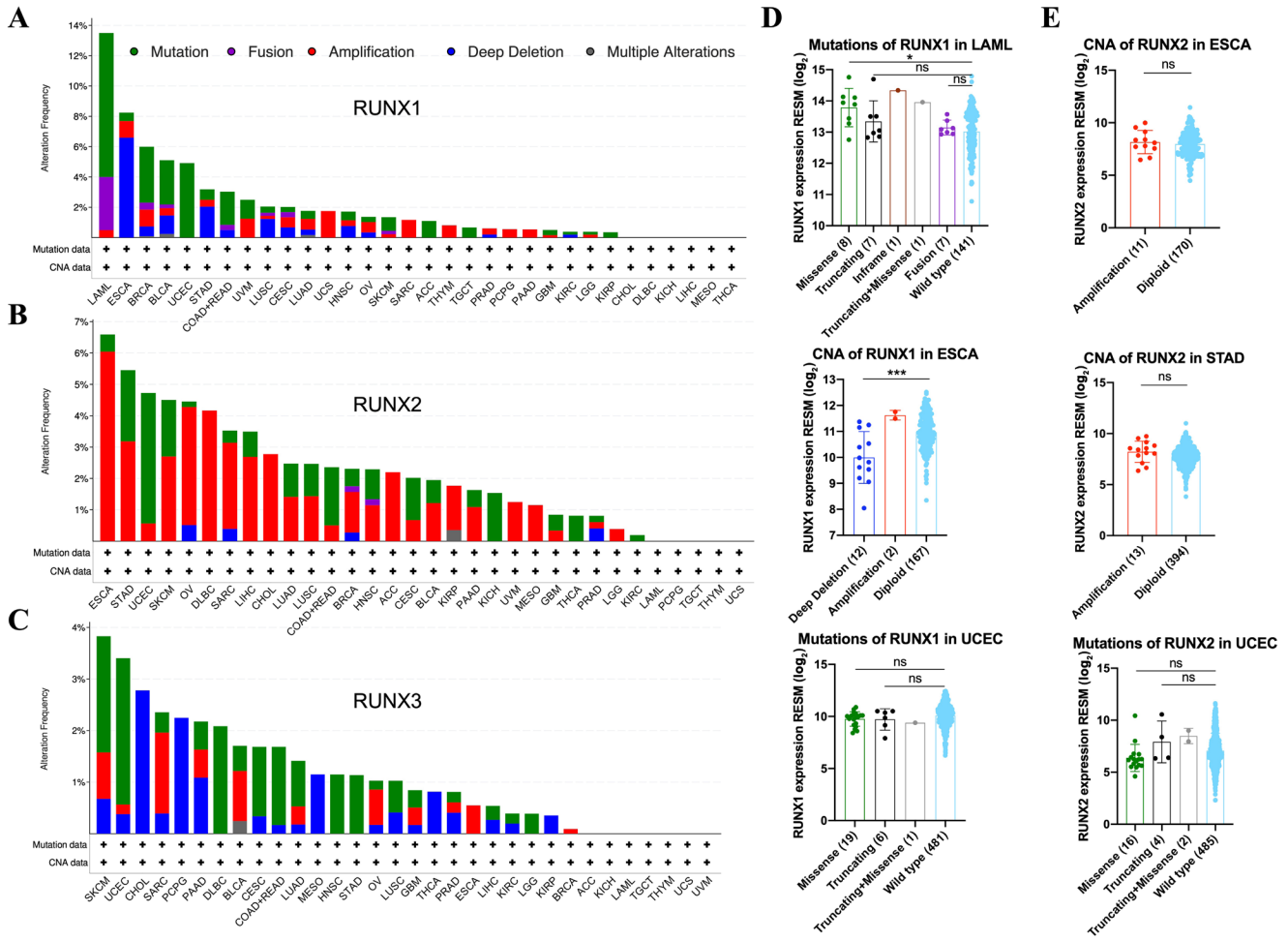

Figure 4: Pan-cancer genetic alterations in *RUNX* genes. Bar graphs showing the frequencies of genetic alteration in (A) *RUNX1*, (B) *RUNX2*, and (C) *RUNX3* genes. Relationship between the genetic alteration and expression of (D) *RUNX1* in LAML, ESCA, and UCEC and (E) *RUNX2* in ESCA, STAD, and UCEC. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. LAML: acute myeloid leukemia; ESCA: esophageal carcinoma; UCEC: uterine corpus endometrial carcinoma; STAD: stomach adenocarcinoma; ns: not significant ($P > 0.05$); BRCA: breast carcinoma; BLCA: bladder urothelial carcinoma; COAD: colon adenocarcinoma; UVM: uveal melanoma; LUSC: lung squamous cell carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; LUAD: lung adenocarcinoma; HNSC: head and neck squamous cell carcinoma; SKCM: skin cutaneous melanoma; THYM: thymoma; TGCT: testicular germ cell tumors; PRAD: prostate adenocarcinoma; PAAD: pancreatic adenocarcinoma; GBM: glioblastoma multiforme; KIRC: kidney renal clear cell carcinoma; LGG: lower-grade glioma; KIRP: kidney renal papillary cell carcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; KICH: kidney chromophobe; LIHC: liver hepatocellular carcinoma; THCA: thyroid carcinoma; READ: rectum adenocarcinoma; UCS: uterine carcinosarcoma; OV: ovarian serous cystadenocarcinoma; SARC: sarcoma; ACC: adrenocortical carcinoma; PCPG: pheochromocytoma and paraganglioma; CHOL: cholangiocarcinoma; MESO: mesothelioma.

had nine frameshift mutations, including eight insertions and one deletion among eight BRCA cases and one LAML case (Figure 5A). The hotspot in *RUNX2* was S31Pfs*9 comprising five frameshift deletion samples, including four COAD patients and one STAD patient (Figure 5B). In *RUNX3*, which had the lowest mutation frequency, a missense mutation of the hotspot R205C was found in two COAD patients and one GBM patient (Figure 5A). The co-occurrence of mutations was primarily observed in patients with *RUNX2* mutations. Only one COAD patient exhibited mutations in *RUNX1, RUNX2*, and *RUNX3*. Meanwhile, 12 patients with UCEC exhibited *RUNX* family co-mutations, which was the most frequent among all cancers (Figure 5B).

The distribution of mutation sites in the protein domain revealed significantly more mutations in the Runt domain than in the RunxI domain in the RUNX family proteins. Specifically, the mutation rate in the Runx1 domain exceeded 55%, while that in the RunxI domain was <5% (Figure 5C). Although a higher correlation was observed in individual cancers, the expression of RUNX family genes was consistently negatively correlated only with TMB in ESCA, KIRP, LIHC, PRAD, and UCEC. The correlation between BRCA and LAML with RUNX1 was approximately −0.4 and 0.4, respectively. Meanwhile, the TMB of kidney chromophobe (KICH) was only significantly correlated (0.4) with *RUNX2*, whereas *RUNX3* was associated with the least tumors significantly

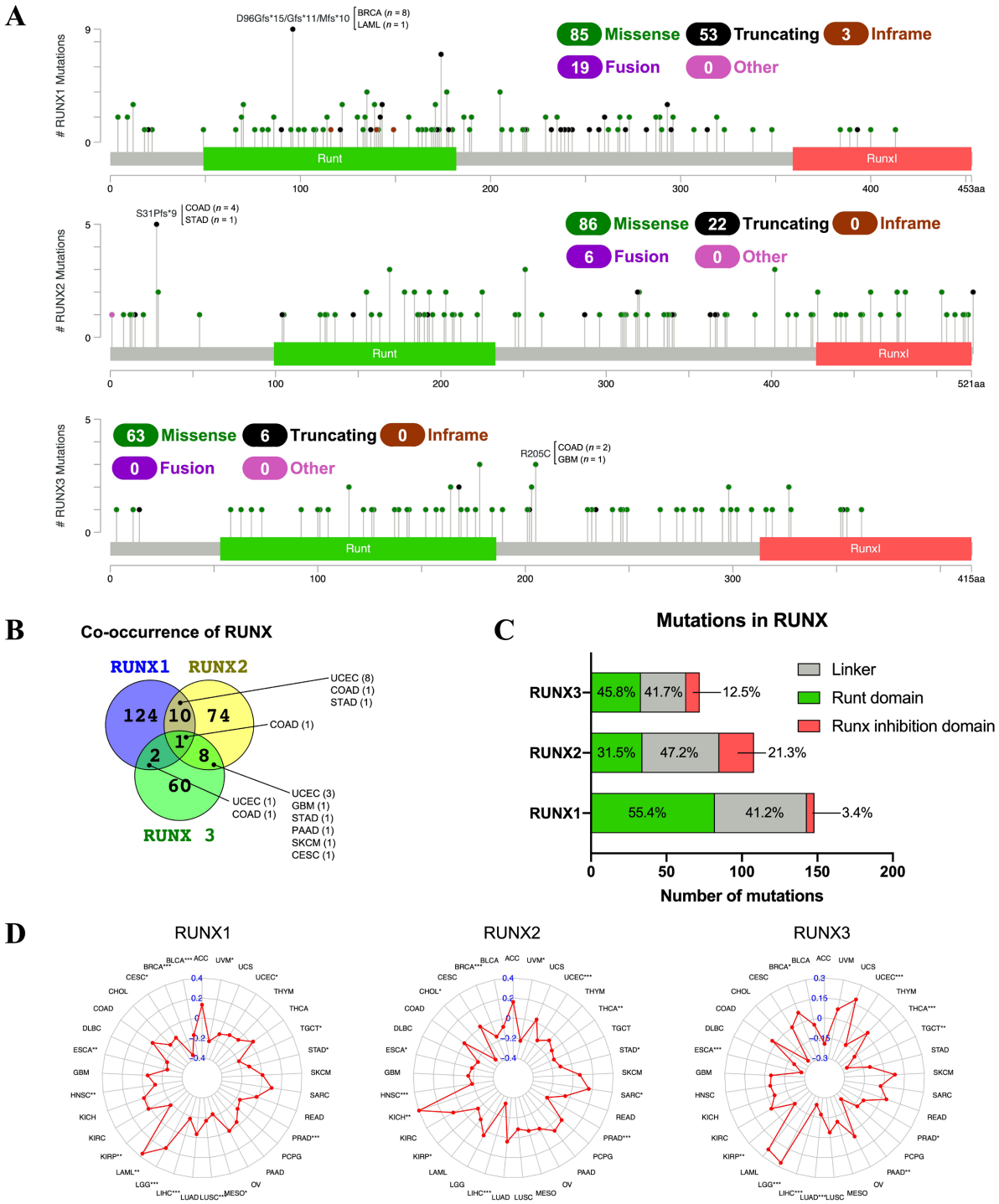

Figure 5: *RUNX* mutations in cancer. (A) Pan-cancer amino acid mutations in *RUNX* in TCGA. Hotspot mutations are indicated. (B) Co-occurrence of *RUNX* mutations. (C) The distribution frequency of mutations in the domains of each RUNX family protein. (D) Correlation between the expression of *RUNX* genes and TMB. **P* < 0.05, ***P* < 0.01, and ****P* < 0.001. TCGA: The Cancer Genome Atlas; TMB: tumor mutation burden; COAD: colon adenocarcinoma; STAD: stomach adenocarcinoma; GBM: glioblastoma multiforme; UCEC: uterine corpus endometrial carcinoma; PAAD: pancreatic adenocarcinoma; SKCM: skin cutaneous melanoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; UVM: uveal melanoma; THYM: thymoma; THCA: thyroid carcinoma; TGCT: testicular germ cell tumors; PRAD: prostate adenocarcinoma; LUSC: lung squamous cell carcinoma; LUAD: lung adenocarcinoma; LIHC: liver hepatocellular carcinoma; LGG: lower-grade glioma; LAML: acute myeloid leukemia; KIRP: kidney renal papillary cell carcinoma; KIRC: kidney renal clear cell carcinoma; KICH: kidney chromophobe; HNSC: head and neck squamous cell carcinoma; GBM: glioblastoma multiforme; ESCA: esophageal carcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; BRCA: breast carcinoma; BLCA: bladder urothelial carcinoma; ACC: adrenocortical carcinoma; UCS: uterine carcinosarcoma; SARC: sarcoma; MESO: mesothelioma; READ: rectum adenocarcinoma; PCPG: pheochromocytoma and paraganglioma; OV: ovarian serous cystadenocarcinoma; CHOL: cholangiocarcinoma.

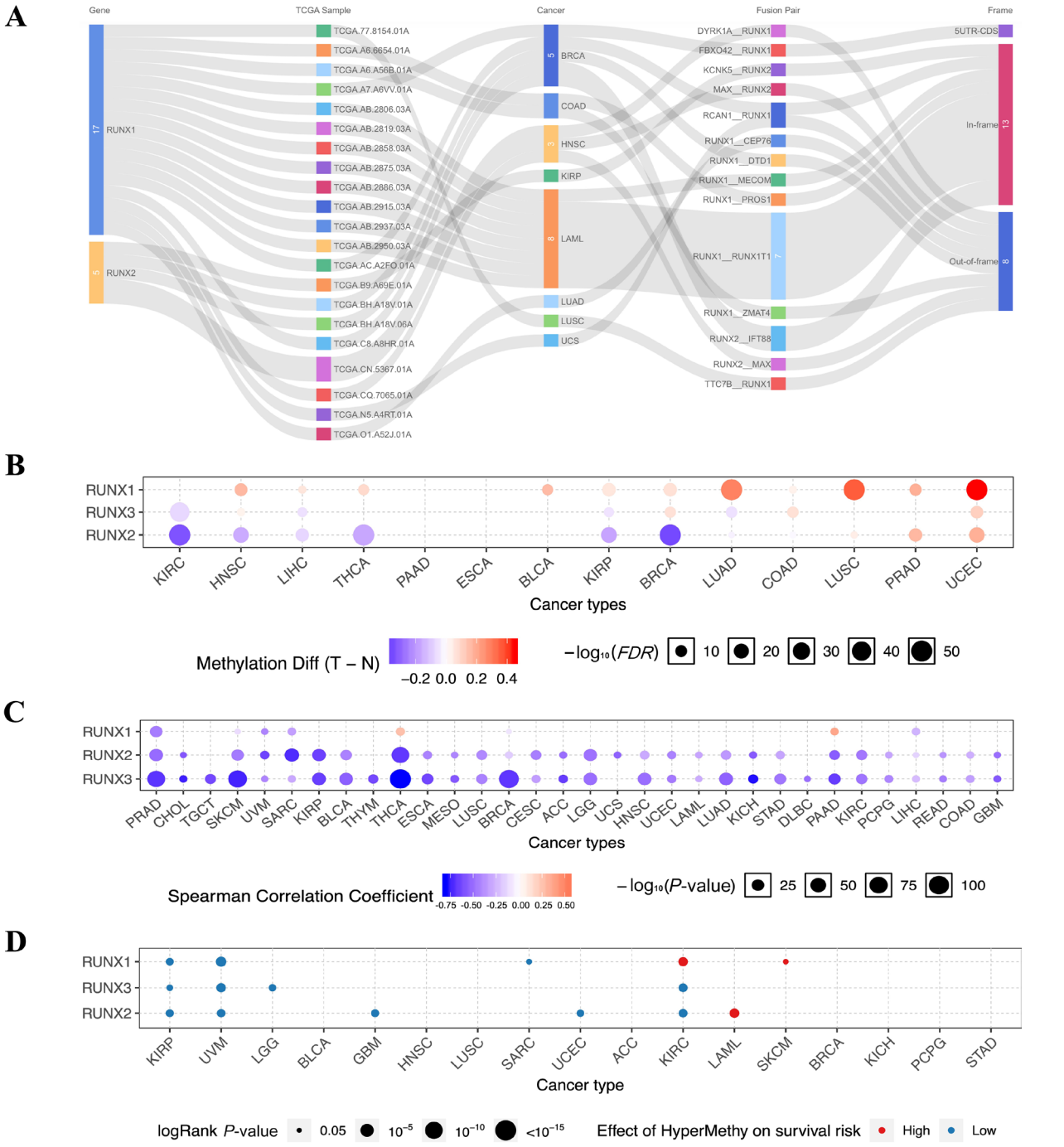

Figure 6: Pan-cancer fusion gene and methylation analysis of *RUNX*. (A) Fusion gene analysis of *RUNX* in cancer. (B) Differential methylation bubble plot showing changes in RUNX methylation between tumor and normal samples for each cancer type. Blue points represent reduced methylation in tumors; red points represent increased methylation in tumors; the deeper the color, the higher the difference; point size represents significance; bigger the point, higher the significance. (C) Pearson correlation between methylation and *RUNX* mRNA expression. Blue points represent negative correlations; red points represent positive correlations; the deeper the color, the greater the correlation. (D) Survival difference between individuals with hypermethylated and hypomethylated *RUNX* genes. Only genes with a significant log *P*-value (≤0.05) are shown in the figure. The red point indicates the high risk of the high-methylation group; the blue point indicates the low risk of the high-methylation group; BRCA: breast carcinoma; COAD: colon adenocarcinoma; HNSC: head and neck squamous cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LAML: acute myeloid leukemia; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; KIRC: kidney renal clear cell carcinoma; LIHC: liver hepatocellular carcinoma; THCA: thyroid carcinoma; PAAD: pancreatic adenocarcinoma; ESCA: esophageal carcinoma; BLCA: bladder urothelial carcinoma; PRAD: prostate adenocarcinoma; UCEC: uterine corpus endometrial carcinoma; TGCT: testicular germ cell tumors; SKCM: skin cutaneous melanoma; UVM: uveal melanoma; THYM: thymoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; LGG: lower-grade glioma; KICH: kidney chromophobe; STAD: stomach adenocarcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; GBM: glioblastoma multiforme; ACC: adrenocortical carcinoma; UCS: uterine carcinosarcoma; SARC: sarcoma; MESO: mesothelioma; READ: rectum adenocarcinoma; PCPG: pheochromocytoma and paraganglioma; CHOL: cholangiocarcinoma.

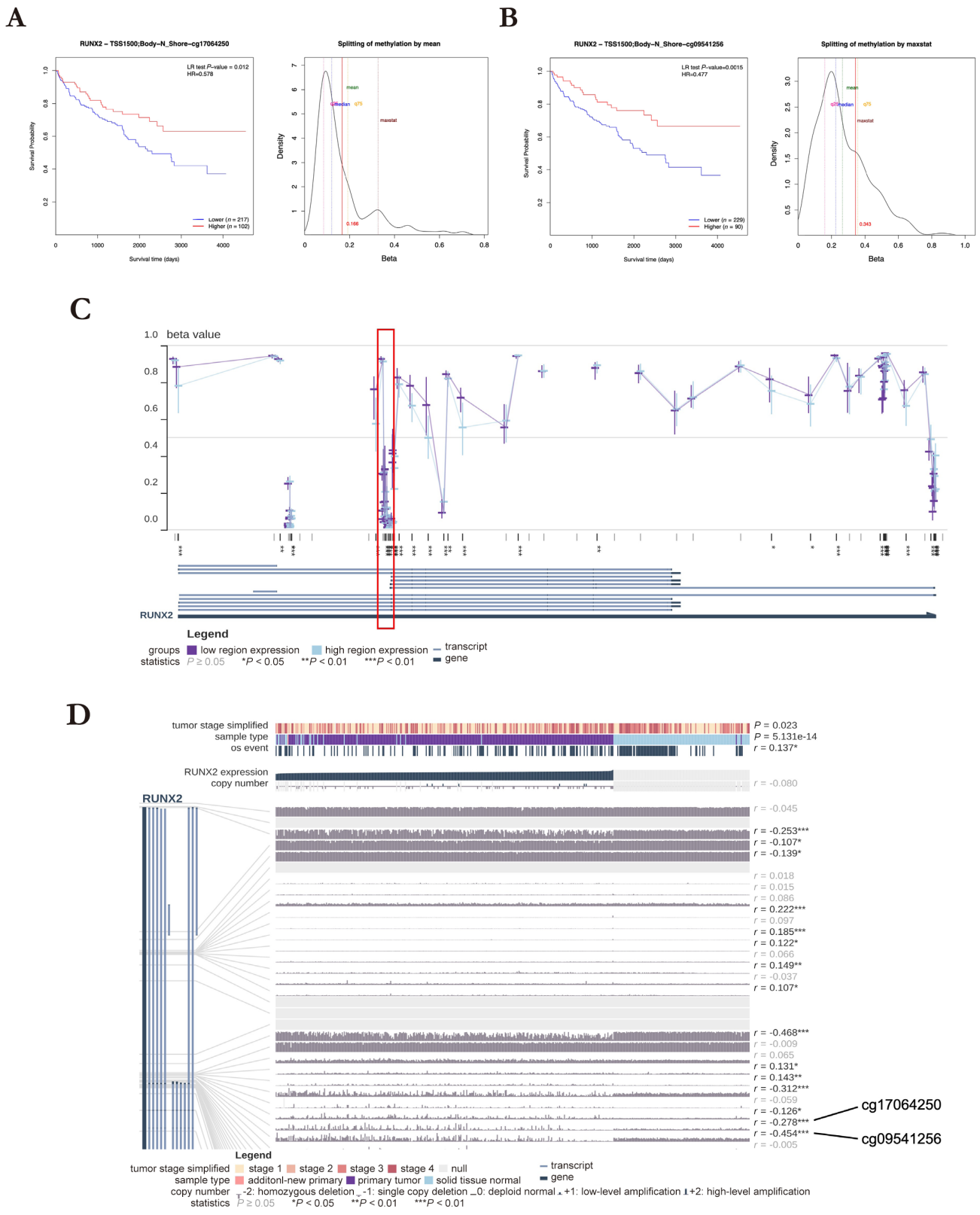

Figure 7: Analysis of the methylation probes, cg17064250 and cg09541256, upstream of the *RUNX2* TSS in KIRC. Survival analysis and distribution of the methylation level of the *RUNX2* probes, (A) cg17064250 and (B) cg09541256, in KIRC. (C) A schematic representing the relationship between the methylation level of all *RUNX2* probes and the expression of *RUNX2*. The red box marked the methylation level in the upstream of RUNX2 TSS. (D) The upstream probes, cg17064250 and cg09541256, of the *RUNX2* partial transcript TSS exhibit a significant negative correlation with *RUNX2* expression. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. TSS: transcription start site; KIRC: kidney renal clear cell carcinoma.

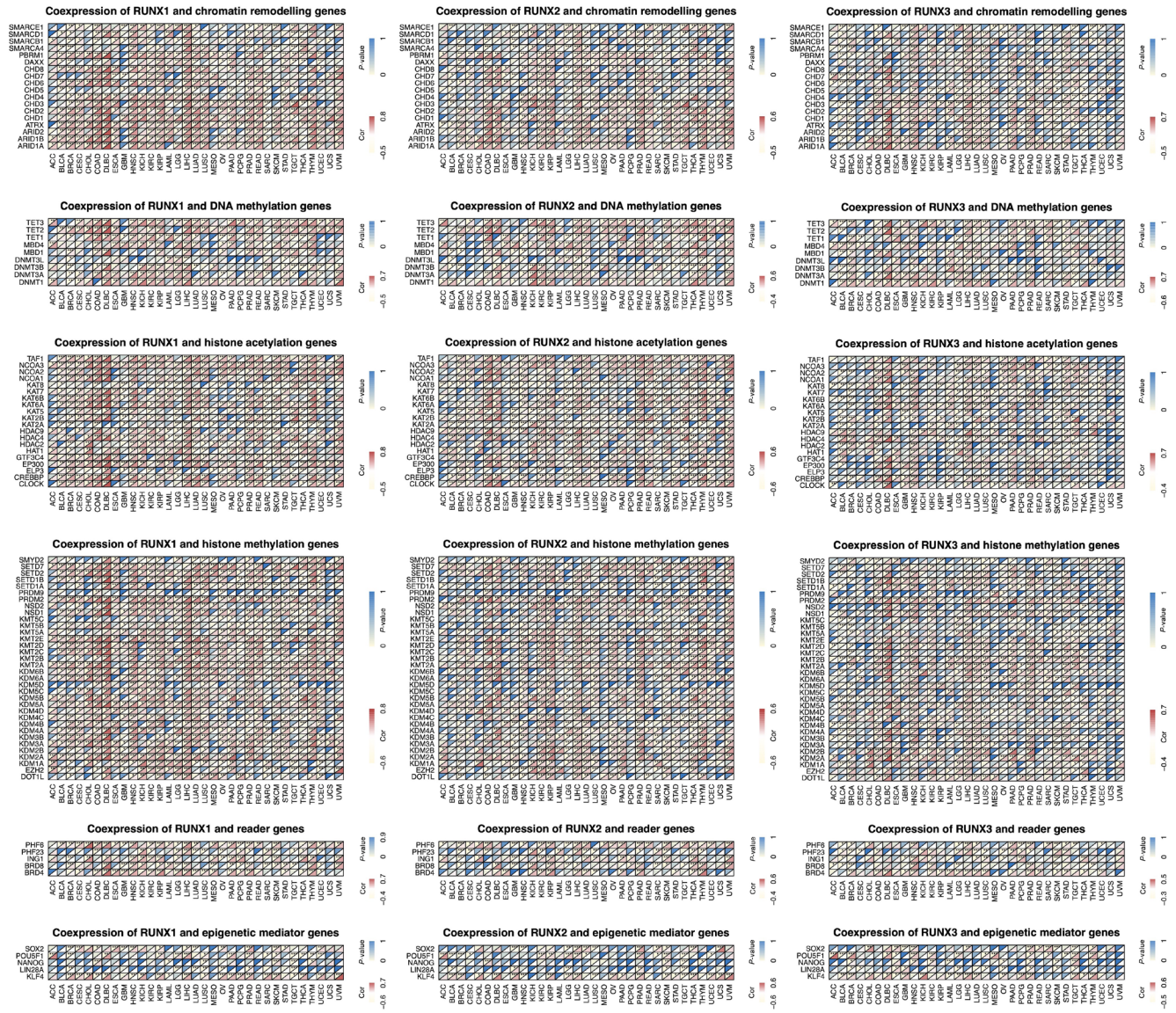

Figure 8: Co-expression of *RUNX* genes and epigenetic regulators. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. BLCA: bladder urothelial carcinoma; BRCA: breast carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD: colon adenocarcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: esophageal carcinoma; GBM: glioblastoma multiforme; HNSC: head and neck squamous cell carcinoma; KICH: kidney chromophobe; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LAML: acute myeloid leukemia; LGG: lower-grade glioma; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; PAAD: pancreatic adenocarcinoma; PRAD: prostate adenocarcinoma; SKCM: skin cutaneous melanoma; STAD: stomach adenocarcinoma; TGCT: testicular germ cell tumors; THCA: thyroid carcinoma; THYM: thymoma; UCEC: uterine corpus endometrial carcinoma; UVM: uveal melanoma; CHOL: cholangiocarcinoma; MESO: mesothelioma; OV: ovarian serous cystadenocarcinoma; PCPG: pheochromocytoma and paraganglioma; READ: rectum adenocarcinoma; SARC: sarcoma; UCS: uterine carcinosarcoma.

correlated with TMB (Figure 5D).

Fusion gene

Gene fusion is closely associated with the occurrence and development of various diseases, especially cancer, and is even the direct cause of certain cancers.[24] Hence, we used the Sankey diagram to identify all fusion genes of the RUNX family in TCGA. A total of 17 cases of *RUNX1* fusion genes, including eight LAML cases, of which seven comprised the RUNX1–RUNX1T1 fusion, which is a high-frequency chromosomal alteration in LAML (Figure 6A) and plays an important role in LAML development.[25]

DNA methylation analysis

We used the online tool GSCALite to conduct a rough methylation analysis. The methylation level of *RUNX* family genes significantly differed between the tumor and normal samples. In particular, the methylation level of

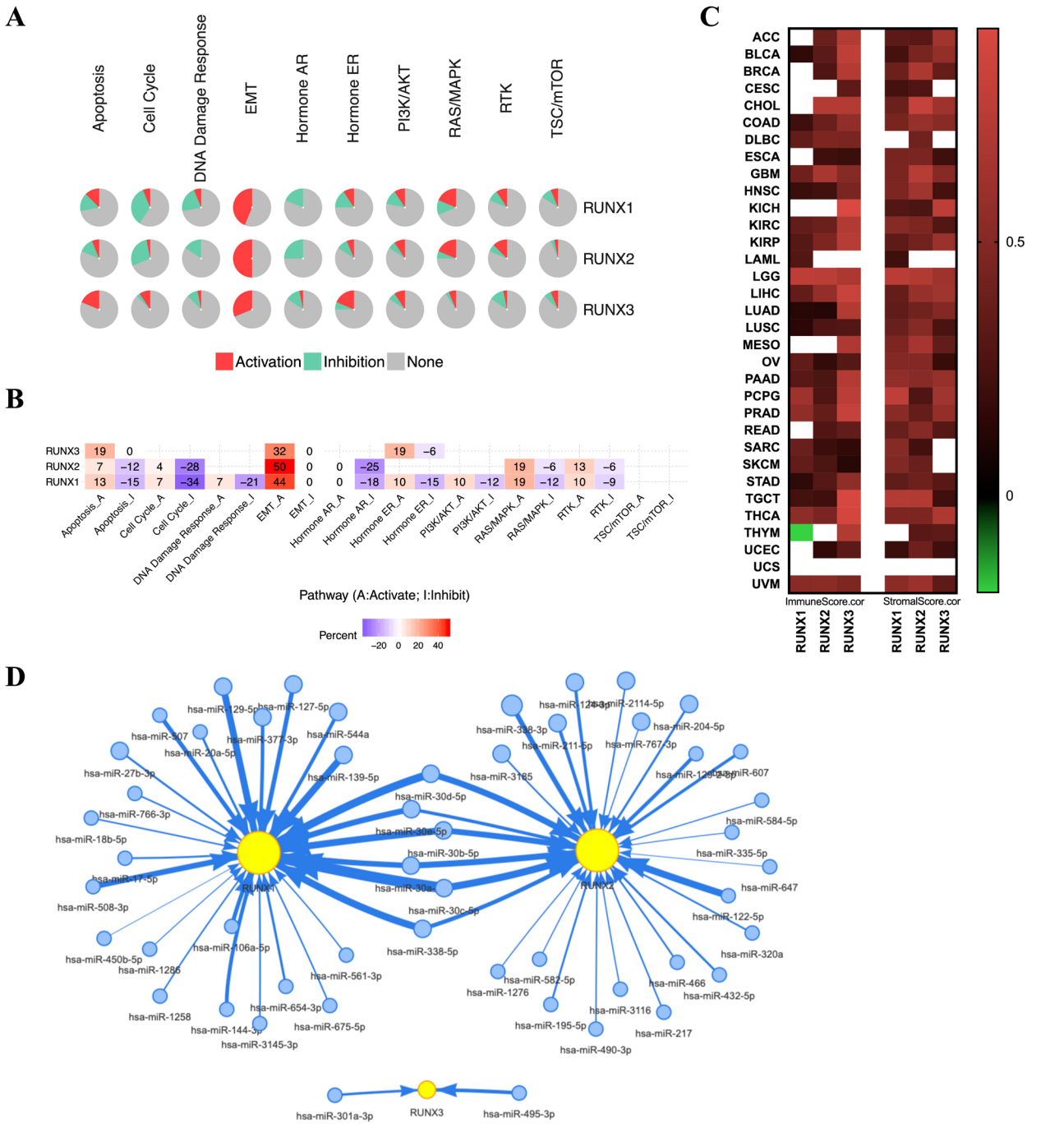

Figure 9: The pan-cancer pathway and miRNA network of *RUNX* genes. (A) Pathway activity showing the overall influence of RUNX on the tumorigenesis pathway in 32 cancer types ([the number of activated or inhibited cancers/32] × 100%). (B) Heatmap showing RUNX proteins that are functional (inhibited or activated) in at least five cancer types. Pathway_A (red) represents the percentage of cancers whose pathways can be activated by a given gene; pathway_I (blue) represents the percentage of cancers whose pathways can be inhibited by a given gene. (C) Correlation between *RUNX* genes and immune and stroma infiltration scores. (D) The pan-cancer miRNA network regulating *RUNX* genes. Edge width is defined by the absolute value of the correlation coefficient. BLCA: bladder urothelial carcinoma; BRCA: breast carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD: colon adenocarcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: esophageal carcinoma; GBM: glioblastoma multiforme; HNSC: head and neck squamous cell carcinoma; KICH: kidney chromophobe; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LAML: acute myeloid leukemia; LGG: lower-grade glioma; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; PAAD: pancreatic adenocarcinoma; PRAD: prostate adenocarcinoma; SKCM: skin cutaneous melanoma; STAD: stomach adenocarcinoma; TGCT: testicular germ cell tumors; THCA: thyroid carcinoma; THYM: thymoma; UCEC: uterine corpus endometrial carcinoma; UVM: uveal melanoma; CHOL: cholangiocarcinoma; MESO: mesothelioma; OV: ovarian serous cystadenocarcinoma; PCPG: pheochromocytoma and paraganglioma; READ: rectum adenocarcinoma; SARC: sarcoma; UCS: uterine carcinosarcoma.

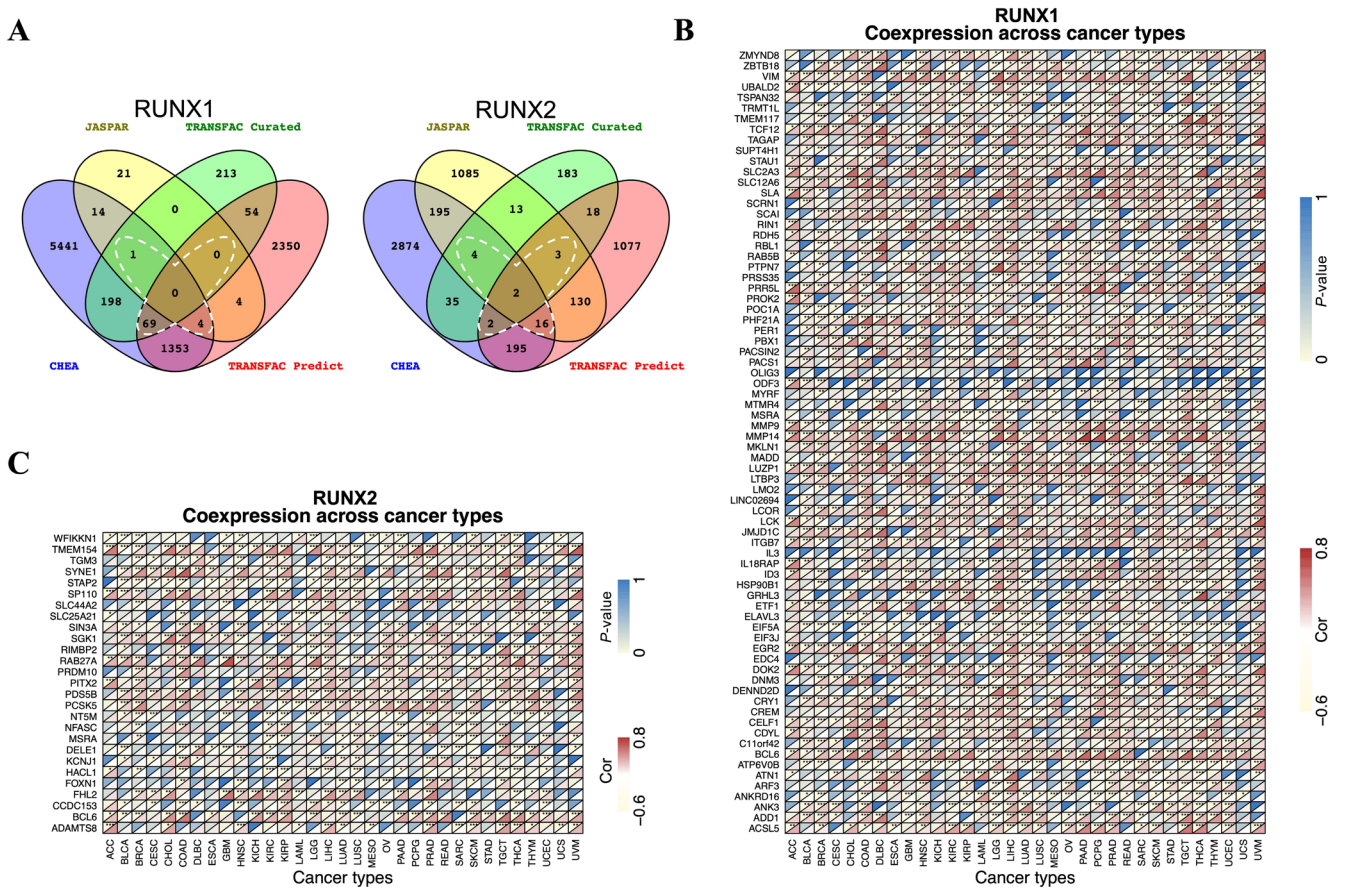

Figure 10: The transcriptional regulatory genes of RUNX proteins. (A) Genes appearing in at least three target gene sets are defined as candidate targets of *RUNX* genes. The white dashed line delineates candidate target genes in the Venn diagram. Pan-cancer heatmap showing the correlation between candidate target genes and (B) *RUNX1* and (C) *RUNX2*. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. BLCA: bladder urothelial carcinoma; BRCA: breast carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD: colon adenocarcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: esophageal carcinoma; GBM: glioblastoma multiforme; HNSC: head and neck squamous cell carcinoma; KICH: kidney chromophobe; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LAML: acute myeloid leukemia; LGG: lower-grade glioma; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; PAAD: pancreatic adenocarcinoma; PRAD: prostate adenocarcinoma; SKCM: skin cutaneous melanoma; STAD: stomach adenocarcinoma; TGCT: testicular germ cell tumors; THCA: thyroid carcinoma; THYM: thymoma; UCEC: uterine corpus endometrial carcinoma; UVM: uveal melanoma.

RUNX1 increased significantly in LUAD, lung squamous cell carcinoma (LUSC), and UCEC. Indeed, lung cancer epigenetic changes are important factors in its occurrence and development.[26,27] Meanwhile, *RUNX2* methylation was significantly reduced in KIRC and BRCA (Figure 6B). Moreover, a significant and negative correlation was detected between methylation and *RUNX2* and *RUNX3* expression (Figure 6C). Additionally, survival risk analysis revealed a consistent trend of high methylation and low-risk *RUNX* family genes in KIRP and UVM (Figure 6D).

Our results, combined with those of previous transcriptome analyses, revealed that *RUNX2* was overexpressed in KIRC (Figure 1D), possibly due to decreased methylation, which may lead to tumorigenesis and poor prognosis (Figure 6B, D). Moreover, according to the "All cancers" module in MethSurv, which provides the results of Cox proportional-hazards analysis, among the methylation probes near the TSS of *RUNX2* in KIRC, cg09541256 and cg17064250 – grouped according to the best split – showed a poor prognosis in the hypomethylation groups (Figure 7A, B). Meanwhile, MEXPRESS further revealed that of the 480 KIRC samples, some transcripts with high *RUNX2* expression had lower methylation upstream of TSS (Figure 7C). Moreover, a significant negative correlation between cg09541256 and cg17064250 methylation and *RUNX2* mRNA expression was observed (Figure 7D).

RUNX and its epigenetic regulators

RUNX genes exhibited significant positive correlations with many epigenetic regulators, suggesting that *RUNX* genes and epigenetic regulators interact with each other or have a regulatory relationship. Among the *RUNX* genes, *RUNX1*

had the highest correlation with epigenetic modifiers. In lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), *RUNX* genes and epigenetic modifiers were more highly correlated than in other malignant tumors. Moreover, most of these epigenetic modifiers were found related to transcription activation (Figure 8). The co-expression of *RUNX* genes and epigenetic regulation provide a basis for further research on specific tumors.

RUNX gene-related cancer pathways

Following the pan-cancer pathway analysis of *RUNX* family genes, the global percentage data revealed that the EMT pathway was activated, not inhibited (Figure 9A). Moreover, the heatmap revealed that 50% of the cancers have EMT related to *RUNX2* (Figure 9B).

ESTIMATE scored the pan-cancer stroma and immune invasion levels based on TCGA transcript data, while Pearson correlation analysis was performed using the obtained score and the expression level of the *RUNX* family genes. The heatmap showed a significant ($P < 0.05$) correlation between the *RUNX* family genes and pan-cancer stroma and immune scores (Figure 9C). That is, the immune and stromal scores in most cancers were positively correlated with *RUNX* gene expression, possibly indicating that RUNX correlates with tumor purity. A major component of the stroma is cancer-associated fibroblasts (CAFs) that mediate EMT.[28]

miRNA network in RUNX

Using GSCALite, we predicted totally 51 miRNAs potentially targeting the 3'-untranslated region (3'-UTR) of *RUNX* transcripts. Pan-cancer analysis revealed that miRNAs potentially inhibit *RUNX1* and *RUNX2* expression more than *RUNX3* expression. Specifically, *RUNX1* and *RUNX2* were both inhibited by six miRNAs, namely, hsa-miR-30a-5p, hsa-miR-30b-5p, hsa-miR-30c-5p, hsa-miR-30d-5p, hsa-miR-30e-5p, and hsa-miR-338-5p (Figure 9D). In osteogenic differentiation, DGCR5 upregulated RUNX2 through miR-30d-5p.[29] Meanwhile, in giant cell bone tumor, miR-30a inhibited osteolysis by targeting RUNX2.[30]

Targeted RUNX genes

In Harmonizome, the four target gene sets selected included putative target genes that had been verified, manually collected, predicted by the TF motif, and enriched based on high-throughput data, such as ChIP-seq. We selected genes contained in three or four gene sets as candidate target genes of the RUNX protein (Figure 10A). Among the RUNX proteins, RUNX3 was excluded from subsequent analysis due to the lack of relevant data. The correlation analysis suggested that 74 *RUNX1* target genes and 27 *RUNX2* target genes were potentially involved in regulatory mechanisms in various cancers (Figure 10B, C). In particular, the correlation between MMP14 and RUNX1 in PAAD reached 0.84 (Figure 10B), suggesting a possible direct regulatory relationship. Moreover, in GBM, a strong correlation was observed between RAB27A and RUNX2, reaching nearly 0.78 (Figure 10C).

DISCUSSION

RUNX proteins are highly conserved TFs that play a role in blood and blood-related cell lineages during both the developmental and adult life stages.[31] Abnormally expressed *RUNX* genes have been observed in various cancers and have been shown to play a key role in carcinogenesis in some, while eliciting a tumor-suppressor effect in other cancer types.[5] Our research focused on the pan-cancer abnormal changes in *RUNX* genes, revealing that abnormal *RUNX* expression is associated with the prognosis of a variety of cancers. In particular, *RUNX1* was overexpressed in cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), COAD, GBM, KIRC, and KIRP, with high expression found to be associated with a worse prognosis. In contrast, *RUNX2* overexpression is only associated with a poor prognosis in bladder urothelial carcinoma (BLCA) and KIRC. Meanwhile, *RUNX3* exhibited a paradoxical association with prognosis, that is, its downregulation was associated with poor prognosis in BRCA and LIHC, while its upregulation was associated with poor prognosis in LAML and SKCM. Moreover, results from the total protein analysis in the six cancers were inconsistent with the transcriptome data, with RUNX3 exhibiting an opposite trend in LUAD. Additionally, comprehensive analysis of protein phosphorylation, genetic alteration, DNA methylation, co-expression with epigenetic regulators, and miRNA networks revealed the potential carcinogenic mechanism of *RUNX* genes.

High-throughput data obtained from pan-cancer gene analysis has revealed that TFs with powerful biological functions were hot genes. For instance, *MYC* is a unique oncogene in human cancers and its alterations are mutually exclusive to PIK3CA, PTEN, APC, or BRAF.[32] Moreover, NFE2L2 has been identified as a potential pan-cancer prognostic biomarker that is associated with immune infiltration.[33] Similarly, TFAP4 is abnormally expressed in most malignancies and is closely associated with OS as well as the degree of tumor infiltration.[34] While these studies were all performed with a single TF target, the current study assessed three TFs of the RUNX family. Moreover, we have carried out in-depth analysis of proteins, DNA methylation, and so on. Finally, we verified that *RUNX2*

is significantly overexpressed and associated with poor prognosis in KIRC and BLCA. These findings agree with those of previous studies that reported RUNX2 as being associated with cancer progression.[35,36]

TCGA contains data for 33 types of cancer, which can be further divided based on the source of the organ system or histopathologic classification.[37] Pan-squamous cell carcinoma (PSCC) arises from the epithelia of the aerodigestive and genitourinary tracts and includes LUSC, HNSC, ESCA, CESC, and BLCA.[38] RUNX mRNAs were independently overexpressed in four of the five PSCCs, while ESCA and HNSC were overexpressed in all. Moreover, *RUNX3* was downregulated in five cancers, including BRCA, COAD, LIHC, LUAD, and thyroid carcinoma (THCA), which are all adenocarcinomas or found in organs with secretory functions. Meanwhile, in PRAD, both *RUNX1* and *RUNX2* were downregulated. Therefore, identifying these similarities and differences in *RUNX* gene expression within different cancers – originating from different tissues and with different pathological characteristics – may pave the way for future studies on *RUNX* genes.

Phosphorylation is a ubiquitous posttranslational modification that, in cancer, can regulate cell proliferation, differentiation, apoptosis, and other functions. In fact, specific phosphorylation sites are recognized as potential therapeutic targets.[39–41] Most phosphorylation sites reported in *RUNX1*, including the S21 residue, are modified to activate transcriptional activity.[42] Additionally, we found increased phosphorylation levels of RUNX1 (T14, T18, and S212) in a variety of cancers. For instance, RUNX2 S319 phosphorylation – a potential diagnostic and therapeutic target – was closely associated with the occurrence and progression of PRAD.[43] Moreover, S28 and S340 residues in RUNX2 are potential therapeutic targets as they have previously been shown to regulate the transcriptional activity of *RUNX2*.[44] Phosphorylation can regulate the ubiquitination and degradation of proteins;[45,46] hence, the preservation and activation of RUNX protein functions may also be related to this mechanism.

Point mutations in *RUNX1* and RUNX1–RUNX1T1 fusions are common and critical pathogenic mechanisms in LAML and are related to patient prognosis.[47] Our analysis of genetic alterations also revealed that *RUNX1* mutation could result in increased *RUNX1* expression. Moreover, *RUNX1* expression was most related to the TMB of LAML pan-cancer, revealing a potential mechanism underlying RUNX1-mediated induction of LAML. In fact, all *RUNX* genes were overexpressed in LAML; however, survival analysis showed that only the high *RUNX3* expression group had a worse prognosis in LAML. Interestingly, in LGG, *RUNX* genes showed a consistent trend in survival analysis (OS and DFS), that is, overexpression was associated with poor prognosis. However, *RUNX* mRNA was not overexpressed in LGG compared with that in healthy tissues. Therefore, the prognostic value of *RUNX* genes in LGG requires further exploration. *RUNX1* genetic alteration in BRCA was the third most common among the 33 cancers, and it primarily comprised mutations. Eight of the nine truncating mutations at the D96 hotspot could be tagged to BRCA. Moreover, 12 cases of co-occurrence mutations were detected in *RUNX* genes in UCEC, the most frequent among all cancers. Moreover, in UCEC, the expression of all *RUNX* genes negatively correlated with TMB. There are a few studies on *RUNX* genes in UCEC, and these results may aid future research and treatment.

As dominant oncogenes or tumor suppressors, *RUNX* genes playing dual roles in cancer have garnered considerable research attention.[48] *RUNX1* has been shown to inhibit cancer stem cell and tumor growth in BRCA, while also acting as a suppressor of EMT in normal and early-stage breast cancer cells. Meanwhile, *RUNX2* has the opposite effect and is involved in a regulatory network that controls EMT.[49,50] This is consistent with the poor prognosis of the *RUNX1* low-expression group, as well as the observed upregulation of *RUNX2* mRNA and protein expression in BRCA. Approximately 1 million new cases of gastric cancer are reported worldwide each year.[51] *RUNX2* has been reported as overexpressed in gastric cancer tissues and promotes metastasis by upregulating CXCR4.[52] In contrast, *RUNX3* acts as a tumor suppressor in gastric cancer to inhibit tumor proliferation and metastasis.[53,54]

Most cancer-related reports on DNA methylation of *RUNX* genes are based on *RUNX3* studies. DNA hypermethylation – responsible for a decrease in *RUNX3* expression – is associated with lung cancer.[55] Hypermethylation in the promoter region may be an important cause of *RUNX3* downregulation in ESCA.[56] Moreover, in breast cancer, hypermethylation of the *RUNX3* promoter is considered a valuable marker.[57] Besides, DNA methylation of *RUNX3* has also been reported in gastric cancer,[58] colorectal cancer,[59] and laryngeal squamous cell carcinoma.[60] This finding is also consistent with the results of our study in that the mRNA expression of *RUNX3* negatively correlated with the DNA methylation level in most cancers. TFs themselves can be modified by methylation, a phenomenon that in turn regulates the expression of other genes. Meanwhile, TFs can bind to the distal enhancer, thereby enabling their demethylation. RUNX1 may function as a scaffold that binds to the enhancer elements in KIRC and GBM to regulate the expression of multiple genes.[61] However, in this study, our DNA

methylation analysis – using GSCALite – did not involve the analysis of specific position probes within the gene; therefore, the results can only be presented based on the overall methylation level. Nevertheless, the methylation levels upstream of the *RUNX2* promoter (cg09541256 and cg17064250) identified in our study are not only related to KIRC prognosis, but also to *RUNX2* expression. Therefore, the activities of *RUNX* genes and DNA methylation constitute a potential pan-cancer carcinogenic mechanism.

Our results also suggested that RUNX is involved in co-expression relationships with many epigenetic regulators. Numerous modifiers are significantly related to *RUNX* genes in a variety of malignant tumors. Previous research has supported the interaction of RUNX proteins with epigenetic modifiers, such as histone deacetylases (HDACs),[62,63] the H3K4 methyltransferase in mixed-lineage leukemia (MLL; also known as KMT2A), and acetyltransferase E1A binding protein P300 (EP300).[64–66] RUNX proteins, as TFs, can recruit epigenetic modifiers to activate or repress transcription,[48] and the co-expression relationships with numerous epigenetic regulators revealed the potential coactivators or corepressors of RUNX proteins in terms of epigenetics.

RUNX proteins can mediate EMT through various signaling pathways, and their activity has been confirmed in various cancers.[5,67] Moreover, in our functional analysis, we observed that pan-cancer, RUNX proteins functioned by activating EMT. In our previous study, we also reported that RUNX2 promotes cancer progression by inducing EMT in renal cell carcinoma.[35] In TME, RUNX expression is related to the degree of immune and stromal infiltration in most cancers. This finding also revealed that RUNX proteins have potentially important roles in TME. Moreover, RUNX proteins contribute to the development of cytotoxic T cells and functioning of dendritic cells and macrophages.[68] CAFs, which account for the largest proportion of cells in the matrix, play a role in EMT.[28,69] In BLCA, we verified that *RUNX2* is a CAF infiltration-related gene and significantly promotes cancer progression.[36] Moreover, the observed correlation between *RUNX* genes and the stromal score revealed that *RUNX* genes play a pivotal in CAFs.

We also conducted a comprehensive, multi-dimensional analysis of *RUNX* genes based on biological information. The robust role of *RUNX* genes in cancer is a cause for concern; however, the detailed mechanisms underlying the pan-cancer role of *RUNX* genes and their carcinogenic mechanism warrant further experimental exploration.

CONCLUSION

In summary, this pan-cancer analysis provides a relatively comprehensive understanding of the association of *RUNX* genes with cancers. The abnormal expression of *RUNX* genes causes a disturbance in many cancers and interferes with patient prognosis. Our results indicate that genetic alterations in *RUNX* genes and their co-expression with epigenetic regulators may be the underlying cause of certain cancers. Meanwhile, DNA methylation, miRNA network, and potential target gene analyses revealed the potential mechanisms underlying the abnormal expression and downstream regulation of *RUNX* genes. Hence, the role of *RUNX* genes in cancer warrants further in-depth exploration.

Data Sharing Statement

In results of RUNX protein phosphorylation and DNA methylation analysis, the details of conservation of RUNX protein phosphorylation sites and potential kinases and results of cox proportional-hazard analysis for all methylation probes of *RUNX* genes could be obtained by contacting the corresponding author.

Authors' Contributions

Hou Y and Liu B proposed the study concept, designed the study, and drafted the manuscript. Liu B, Sun S, and Pan S collected, analyzed, and interpreted the data. All authors read and approved the final manuscript.

Source of Funding

The present study was supported by 345 Talent Project from Shengjing Hospital of China Medical University (No. M0312 and M0340), Doctoral Scientific Research Foundation of Liaoning Province (No. 2020-BS-120), and joint plan of research and development program of Liaoning Province (No. 2020JH 2/10300137).

Availability of Data and Materials

The datasets used and/or analyzed during the current study are available from UCSC Xena (https://xenabrowser.net/datapages/).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, *et al*. The Human Transcription Factors. Cell 2018;172:650–65.
2. Papavassiliou KA, Papavassiliou AG. Transcription Factor Drug Targets. J Cell Biochem 2016;117:2693–6.
3. Bushweller JH. Targeting transcription factors in cancer - from undruggable to reality. Nat Rev Cancer 2019;19:611–24.
4. Lambert M, Jambon S, Depauw S, David-Cordonnier MH. Targeting Transcription Factors for Cancer Treatment. Molecules 2018;23:1479.
5. Ito Y, Bae SC, Chuang LS. The RUNX family: developmental regulators in cancer. Nat Rev Cancer 2015;15:81–95.
6. Chuang LS, Ito K, Ito Y. RUNX family: Regulation and diversification of roles through interacting proteins. Int J Cancer 2013;132:1260–71.
7. Kagoshima H, Shigesada K, Satake M, Ito Y, Miyoshi H, Ohki M, *et al*. The Runt domain identifies a new family of heteromeric transcriptional regulators. Trends Genet 1993;9:338–41.
8. Tahirov TH, Inoue-Bungo T, Morii H, Fujikawa A, Sasaki M, Kimura K, *et al*. Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta. Cell 2001;104:755–67.
9. Fukushima-Nakase Y, Naoe Y, Taniuchi I, Hosoi H, Sugimoto T, Okuda T. Shared and distinct roles mediated through C-terminal subdomains of acute myeloid leukemia/Runt-related transcription factor molecules in murine development. Blood 2005;105:4298–307.
10. Sood R, Kamikubo Y, Liu P. Role of RUNX1 in hematological malignancies. Blood 2017;129:2070–82.
11. Villanueva F, Araya H, Briceño P, Varela N, Stevenson A, Jerez S, *et al*. The cancer-related transcription factor RUNX2 modulates expression and secretion of the matricellular protein osteopontin in osteosarcoma cells to promote adhesion to endothelial pulmonary cells and lung metastasis. J Cell Physiol 2019;234:13659–79.
12. Chen F, Chandrashekar DS, Varambally S, Creighton CJ. Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers. Nat Commun 2019;10:5679.
13. 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 2019;47:W556–W560.
14. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, *et al*. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401–04.
15. Liu CJ, Hu FF, Xia MX, Han L, Zhang Q, Guo AY. GSCALite: a web server for gene set cancer analysis. Bioinformatics 2018;34:3771–2.
16. Koch A, Jeschke J, Van Criekinge W, van Engeland M, De Meyer T. MEXPRESS update 2019. Nucleic Acids Res 2019;47:W561–W565.
17. Modhukur V, Iljasenko T, Metsalu T, Lokk K, Laisk-Podar T, Vilo J. MethSurv: a web tool to perform multivariable survival analysis using DNA methylation data. Epigenomics 2018;10:277–88.
18. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nat Rev Genet 2016;17:284–99.
19. Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T, *et al*. New nomenclature for chromatin-modifying enzymes. Cell 2007;131:633–6.
20. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 2012;13:484–92.
21. Akbani R, Ng PK, Werner HM, Shahmoradgoli M, Zhang F, Ju Z, *et al*. A pan-cancer proteomic perspective on The Cancer Genome Atlas. Nat Commun 2014;5:3887.
22. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, *et al*. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013;4:2612.
23. Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, *et al*. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford) 2016;2016.
24. Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. Nat Rev Cancer 2015;15:371–81.
25. Yun JW, Bae YK, Cho SY, Koo H, Kim HJ, Nam DH, *et al*. Elucidation of Novel Therapeutic Targets for Acute Myeloid Leukemias with RUNX1-RUNX1T1 Fusion. Int J Mol Sci 2019;20:1717.
26. Clementsen PF, Bodtger U, Konge L, Christiansen IS, Nessar R, Salih GN, *et al*. Diagnosis and staging of lung cancer with the use of one single echoendoscope in both the trachea and the esophagus: A practical guide. Endosc Ultrasound 2021;10:325–34.
27. Duruisseaux M, Esteller M. Lung cancer epigenetics: From knowledge to applications. Semin Cancer Biol 2018;51:116–28.
28. Fiori ME, Di Franco S, Villanova L, Bianca P, Stassi G, De Maria R. Cancer-associated fibroblasts as abettors of tumor progression at the crossroads of EMT and therapy resistance. Mol Cancer 2019;18:70.
29. Wu ZH, Huang KH, Liu K, Wang GT, Sun Q. DGCR5 induces osteogenic differentiation by up-regulating Runx2 through miR-30d-5p. Biochem Biophys Res Commun 2018;505:426–31.
30. Huang Q, Jiang Z, Meng T, Yin H, Wang J, Wan W, *et al*. MiR-30a inhibits osteolysis by targeting RunX2 in giant cell tumor of bone. Biochem Biophys Res Commun 2014;453:160–5.
31. de Bruijn M, Dzierzak E. Runx transcription factors in the development and function of the definitive hematopoietic system. Blood 2017;129:2061–9.
32. Schaub FX, Dhankani V, Berger AC, Trivedi M, Richardson AB, Shaw R, *et al*. Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas. Cell Syst 2018;6:282–300.e2.
33. Ju Q, Li X, Zhang H, Yan S, Li Y, Zhao Y. NFE2L2 Is a Potential Prognostic Biomarker and Is Correlated with Immune Infiltration in Brain Lower Grade Glioma: A Pan-Cancer Analysis. Oxid Med Cell Longev 2020;2020:3580719.
34. Liu JN, Kong XS, Sun P, Wang R, Li W, Chen QF. An integrated pan-cancer analysis of TFAP4 aberrations and the potential clinical implications for cancer immunity. J Cell Mol Med 2021;25:2082–97.
35. Liu B, Liu J, Yu H, Wang C, Kong C. Transcription factor RUNX2 regulates epithelial-mesenchymal transition and progression in renal cell carcinomas. Oncol Rep 2020;43:609–16.
36. Liu B, Pan S, Liu J, Kong C. Cancer-associated fibroblasts and the related Runt-related transcription factor 2 (RUNX2) promote bladder cancer progression. Gene 2021;775:145451.
37. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, *et al*. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell 2018;173:291–304.e6.
38. Campbell JD, Yau C, Bowlby R, Liu Y, Brennan K, Fan H, *et al*. Genomic, Pathway Network, and Immunologic Features Distinguishing Squamous Carcinomas. Cell Rep 2018;23:194–212.e196.
39. Singh V, Ram M, Kumar R, Prasad R, Roy BK, Singh KK. Phosphorylation: Implications in Cancer. Protein J 2017;36:1–6.
40. Ashton TM, McKenna WG, Kunz-Schughart LA, Higgins GS. Oxidative Phosphorylation as an Emerging Target in Cancer Therapy. Clin Cancer Res 2018;24:2482–90.
41. Marmiroli S, Fabbro D, Miyata Y, Pierobon M, Ruzzene M. Phosphorylation, Signaling, and Cancer: Targets and Targeting. Biomed Res Int 2015;2015:601543.
42. Goyama S, Huang G, Kurokawa M, Mulloy JC. Posttranslational modifications of RUNX1 as potential anticancer targets. Oncogene 2015;34:3483–92.
43. Ge C, Zhao G, Li Y, Li H, Zhao X, Pannone G, *et al*. Role of Runx2 phosphorylation in prostate cancer and association with metastatic disease. Oncogene 2016;35:366–76.
44. Selvamurugan N, Shimizu E, Lee M, Liu T, Li H, Partridge NC. Identification and characterization of Runx2 phosphorylation sites involved in matrix metalloproteinase-13 promoter activation. FEBS Lett 2009;583:1141–6.
45. Filipčík P, Curry JR, Mace PD. When Worlds Collide-Mechanisms at the Interface between Phosphorylation and Ubiquitination. J Mol Biol

2017;429:1097–113.
46. Swatek KN, Komander D. Ubiquitin modifications. Cell Res 2016;26:399–422.
47. Hornung R, Jurinovic V, Batcha AMN, Bamopoulos SA, Rothenberg-Thurley M, Amler S, *et al*. Mediation analysis reveals common mechanisms of RUNX1 point mutations and RUNX1/RUNX1T1 fusions influencing survival of patients with acute myeloid leukemia. Sci Rep 2018;8:11293.
48. Otálora-Otálora BA, Henríquez B, López-Kleine L, Rojas A. RUNX family: Oncogenes or tumor suppressors (Review). Oncol Rep 2019;42:3–19.
49. Hong D, Fritz AJ, Finstad KH, Fitzgerald MP, Weinheimer A, Viens AL, *et al*. Suppression of Breast Cancer Stem Cells and Tumor Growth by the RUNX1 Transcription Factor. Mol Cancer Res 2018;16:1952–64.
50. Fritz AJ, Hong D, Boyd J, Kost J, Finstaad KH, Fitzgerald MP, *et al*. RUNX1 and RUNX2 transcription factors function in opposing roles to regulate breast cancer stem cells. J Cell Physiol 2020;235:7261–72.
51. Hoibian S, Giovannini M, Autret A, Pesenti C, Bories E, Ratone JP, *et al*. Preoperative EUS evaluation of the response to neoadjuvant therapy for gastric and esophagogastric junction cancer is correlated with survival: A single retrospective study of 97 patients. Endosc Ultrasound 2021;10:103–10.
52. Zhou S, Zhang S, Wang L, Huang S, Yuan Y, Yang J, *et al*. BET protein inhibitor JQ1 downregulates chromatin accessibility and suppresses metastasis of gastric cancer via inactivating RUNX2/NID1 signaling. Oncogenesis 2020;9:33.
53. Xue M, Chen LY, Wang WJ, Su TT, Shi LH, Wang L, *et al*. HOTAIR induces the ubiquitination of Runx3 by interacting with Mex3b and enhances the invasion of gastric cancer cells. Gastric Cancer 2018;21:756–64.
54. Yu J, Tian X, Chang J, Liu P, Zhang R. RUNX3 inhibits the proliferation and metastasis of gastric cancer through regulating miR-182/HOXA9. Biomed Pharmacother 2017;96:782–91.
55. Li QL, Kim HR, Kim WJ, Choi JK, Lee YH, Kim HM, *et al*. Transcriptional silencing of the RUNX3 gene by CpG hypermethylation is associated with lung cancer. Biochem Biophys Res Commun 2004;314:223–8.
56. Tonomoto Y, Tachibana M, Dhar DK, Onoda T, Hata K, Ohnuma H, *et al*. Differential expression of RUNX genes in human esophageal squamous cell carcinoma: downregulation of RUNX3 worsens patient prognosis. Oncology 2007;73:346–56.
57. Lau QC, Raja E, Salto-Tellez M, Liu Q, Ito K, Inoue M, *et al*. RUNX3 is frequently inactivated by dual mechanisms of protein mislocalization and promoter hypermethylation in breast cancer. Cancer Res 2006;66:6512–20.
58. Chen W, Gao N, Shen Y, Cen JN. Hypermethylation downregulates Runx3 gene expression and its restoration suppresses gastric epithelial cell growth by inducing p27 and caspase3 in human gastric cancer. J Gastroenterol Hepatol 2010;25:823–31.
59. Mu WP, Wang J, Niu Q, Shi N, Lian HF. Clinical significance and association of RUNX3 hypermethylation frequency with colorectal cancer: a meta-analysis. Onco Targets Ther 2014;7:1237–45.
60. Liu ZH, Liu JJ, Li SS, Yang XM. Association of RUNX3 Methylation with Clinical Outcome and Cell Migration/Invasion in Laryngeal Squamous Cell Carcinoma. Cancer Invest 2016;34:105–13.
61. Yao L, Shen H, Laird PW, Farnham PJ, Berman BP. Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. Genome Biol 2015;16:105.
62. Westendorf JJ, Zaidi SK, Cascino JE, Kahler R, van Wijnen AJ, Lian JB, *et al*. Runx2 (Cbfa1, AML-3) interacts with histone deacetylase 6 and represses the p21(CIP1/WAF1) promoter. Mol Cell Biol 2002;22:7982–92.
63. Guo H, Friedman AD. Phosphorylation of RUNX1 by cyclin-dependent kinase reduces direct interaction with HDAC1 and HDAC3. J Biol Chem 2011;286:208–15.
64. Huang G, Zhao X, Wang L, Elf S, Xu H, Zhao X, *et al*. The ability of MLL to bind RUNX1 and methylate H3K4 at PU.1 regulatory regions is impaired by MDS/AML-associated RUNX1/AML1 mutations. Blood 2011;118:6544–52.
65. Kitabayashi I, Aikawa Y, Nguyen LA, Yokoyama A, Ohki M. Activation of AML1-mediated transcription by MOZ and inhibition by the MOZ-CBP fusion protein. Embo j 2001;20:7184–96.
66. Aikawa Y, Nguyen LA, Isono K, Takakura N, Tagata Y, Schmitz ML, *et al*. Roles of HIPK1 and HIPK2 in AML1- and p300-dependent transcription, hematopoiesis and blood vessel formation. Embo j 2006;25:3955–65.
67. Voon DC, Thiery JP. The Emerging Roles of RUNX Transcription Factors in Epithelial-Mesenchymal Transition. Adv Exp Med Biol 2017;962:471–89.
68. Seo W, Taniuchi I. The Roles of RUNX Family Proteins in Development of Immune Cells. Mol Cells 2020;43:107–13.
69. Hu JL, Wang W, Lan XL, Zeng ZC, Liang YS, Yan YR, *et al*. CAFs secreted exosomes promote metastasis and chemotherapy resistance by enhancing cell stemness and epithelial-mesenchymal transition in colorectal cancer. Mol Cancer 2019;18:91.

How to cite this article: Pan S, Sun S, Liu B, Hou Y. Pan-cancer landscape of the RUNX protein family reveals their potential as carcinogenic biomarkers and the mechanisms underlying their action. J Transl Intern Med 2022; 10: 156-174.