Biomedical Data Science
2025, VOL. 5, NO. 3, 119-124
DOI: 10.47297/wspbdsWSP2752-630524.20250503

A Potential Marker for Neoplastic Proliferation and Treatment Response in Colorectal Cancer

Mingxuan Li

Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, M13 9PL, United Kingdom

ABSTRACT
Colorectal cancer (CRC) is a prevalent malignancy characterized by high mortality, unclear pathogenesis, and a lack of effective diagnostic and prognostic biomarkers. PI4K2B, a key member of the PI4K family, plays a crucial role in synthesizing the second messenger inositol triphosphate (IP3). However, research on the relationship between PI4K2B and cancer remains limited, and its role in CRC is unclear. This study leverages multiple public databases for bioinformatics analyses to elucidate the correlation between PI4K2B and CRC, revealing its potential dual role as both a driver of CRC development and a protective factor for CRC prognosis.

KEYWORDS
Biomarker; Colorectal cancer; PI4K family

1 Introduction

CRC is a highly prevalent and deadly cancer, accounting for approximately 10% of annual cancer diagnoses and cancer-related deaths worldwide and ranking as the third leading cause of cancer mortality globally. However, the mechanisms underlying CRC remain elusive, and the search for effective diagnostic and prognostic biomarkers continues.

Studies have revealed a correlation between the phosphatidylinositol-4-kinase (PI4K) family and tumor development and progression. The production of PI4P relies on the activity of four phosphatidylinositol 4-kinases: PI4KIIIa (PI4KA), PI4KIIIb (PI4KB), PI4KIIa (PI4K2A), and PI4KIIb (PI4K2B). PI4KA expression in tumors correlates with overall survival and contributes to an immunosuppressive bone tumor microenvironment by preferentially enriching non-activated and immunosuppressive macrophage populations. The PI4KB gene, located on chromosome 1q, exhibits structural and expression abnormalities in breast cancer.

However, systematic and comprehensive research on PI4K2B in tumors remains lacking. In this study, we conducted bioinformatics analyses based on multiple public databases to investigate the correlation between PI4K2B and CRC, assess its potential as a diagnostic and prognostic marker, and analyze its dual role as both a driver of CRC development and a protective factor for CRC prognosis.

2 Methods

2.1 Differential expression of PI4K2B in tumor and normal tissues

STAR-counts data and clinical information for CRC were downloaded from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.com). We extracted data in TPM format, normalized it using log2(TPM+1), and retained samples with both RNAseq data and clinical information, ultimately utilizing 620 samples for subsequent analyses. Genotype-Tissue Expression (GTEx) data were obtained from version V8 (https://gtexportal.org/home/datasets). The "ggplot2" package in R software (v4.0.3) was used to analyze the expression of PI4K2B in tumor and normal tissues and perform Wilcoxon rank-sum tests. Statistical significance was set at $p < 0.05$. Immunohistochemical images of PI4K2B were acquired from The Human Protein Atlas (HPA) database (https://www.proteinatlas.org/).

2.2 PI4K2B expression in various cell types within tumor tissues

We selected the EMTAB8107 dataset for CRC in the Tumor Immune Single-cell Hub 2 (TISCH2) database (http://tisch.comp-genomics.org), collecting data on 23,176 colorectal cancer cells and cell annotation results.

2.3 Relationship between PI4K2B and CRC Clinical Parameters

The "ggplot2" package in R software was used to analyze PI4K2B expression in relation to various clinical characteristics, including sex, age, and stage. The "ggalluvial" package in R software was used to create Sankey diagrams depicting the relationship between PI4K2B expression and CRC clinical parameters. Kaplan-Meier survival curves for high and low PI4K2B expression groups were generated using the "survival" and "survminer" packages in R software, with intergroup comparisons conducted using the log-rank test.

2.4 CRC prognostic model based on PI4K2B and clinical parameters

Forest plots were generated using the "forestplot" package in R software to display each variable (p-value, hazard ratio

[HR], and 95% confidence interval [CI]). Nomograms were created using the "rms" package in R software to predict 1-year, 3-year, and 5-year OS rates. The nomograms provided a graphical representation of the combined impact of these factors, allowing for individualized prognostic risk assessment based on points associated with each risk factor.

2.5 PI4K2B protein interaction network, Gene Set Enrichment Analysis (GSEA), and gene correlation diagram

The STRING database (http://string-db.org) was used to obtain genes related to PI4K2B for constructing a protein-protein interaction (PPI) network. GSEA was performed on 620 CRC samples from TCGA using the Hallmark gene set (https://www.gsea-msigdb.org/gsea/index.jsp). Spearman correlation coefficients between PI4K2B expression and the expression of other genes in CRC samples were calculated using the "ggstatsplot" package in R software, and gene correlation diagrams were generated.

2.6 Differential gene expression and enrichment analysis

The "Limma" package in R software was used to investigate differential gene expression between high and low PI4K2B expression groups. Thresholds for differentially expressed genes (DEGs) were set at $p < 0.05$ and log2(fold change [FC]) > 1.5 or log2(FC) < -1.5. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using the "ClusterProfiler" package in R software. Univariate Cox regression analysis was employed to identify genes associated with CRC prognosis among the DEGs. Risk scores for 620 CRC samples were calculated based on the expression of these genes using the "ggrisk" package in R software. Kaplan-Meier survival curves were generated using the "survival" and "survminer" packages in R software. A multivariate Cox regression model was constructed, and a receiver operating characteristic (ROC) curve was plotted using the "timeROC" package in R software to evaluate model performance.

2.7 Immune feature analysis

The TIMER algorithm in the "immunedeconv" package in R software was used to calculate immune scores for high and low PI4K2B expression groups, assessing the relative abundance of different immune cell types. SIGLEC15, TIGIT, CD274, HAVCR2, PDCD1, CTLA4, LAG3, and PDCD1LG2 are genes associated with immune checkpoints. The TIMER algorithm in the "immunedeconv" package was used to analyze the correlation between the PI4K2B gene, immune cells, and tumor immunophenotype (TIP), as well as the correlation among immune cells themselves. The "ggClusterNet" package in R software was used for analysis and visualization. Spearman correlation diagrams for PI4K2B with six immune cell types were generated using the TIMER algorithm in the "immunedeconv" package and the "ggstatsplot" package. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was used to predict the immunotherapeutic response. Visualizations were created using the "ggplot2" and "ggpubr" packages in R software.

2.8 Relationship between PI4K2B and drug sensitivity analysis

Preprocessed data from the Genomics of Drug Sensitivity in Cancer (GDSC) website (www.cancerrxgene.org) encompassing drug sensitivity, gene expression, and mutation calls were used to predict treatment response for each sample. The prediction process was implemented using the pRRophetic R package. Half-maximal inhibitory concentration (IC50) values for each sample were estimated using ridge regression. All parameters were set to default values, with batch effects using "combat" and tissue type set to "all." Duplicate gene expression values were summarized as the mean. Spearman correlation diagrams and boxplots were generated using the "ggstatsplot" and "ggplot2" packages in R software.

3 Results

3.1 Expression of PI4K2B in normal tissues, CRC tumor tissues, and various cell types within tumor tissues

Initially, we investigated PI4K2B expression in normal tissues, CRC tumor tissues, and different cell types within tumor tissues. Immunohistochemical staining from the HPA database revealed aberrant upregulation of PI4K2B protein expression in CRC tissues (Figure 1A). PI4K2B expression was significantly upregulated ($p < 0.01$) in 620 tumor samples from the TCGA database compared to adjacent normal samples from TCGA and 779 normal samples from the GTEx database (Figure 1B). Single-cell transcriptome data from TISCH, after cell annotation, yielded a total of 13 cell types: B cells, CD4+ T cells, CD8+ T cells, endothelial cells, enteric glial cells, epithelial cells, fibroblasts, M1 macrophages, tumor cells, mast cells, monocytes, myofibroblasts, and plasma cells (Figure 1C). PI4K2B exhibited the highest average expression in tumor cells and epithelial cells (Figure 1D, E).

3.2 Relationship between PI4K2B and CRC clinical parameters

Next, we explored the association between PI4K2B expression and sex, age, TNM stage, and survival in CRC patients. No significant relationship was observed between PI4K2B expression and sex or age (Figure 2A, B), while PI4K2B expression was significantly higher in stages I-II compared to stages III-IV ($p < 0.01$) (Figure 2C). We also generated Sankey

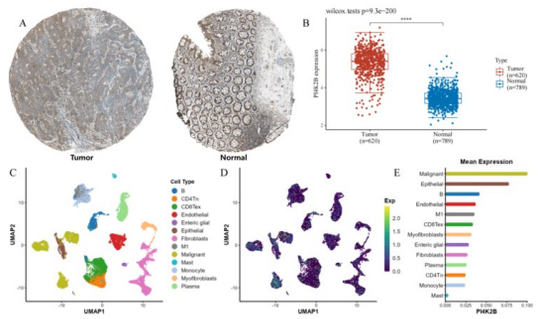

Figure 1 (A) Immunohistochemical staining of PI4K2B in CRC tumor tissues and normal tissues. (B) Expression of PI4K2B in normal and CRC tumor samples. (C) UMAP plot of cell clusters from single-cell RNA sequencing of CRC. (D) UMAP plot of PI4K2B expression in CRC single-cell samples. (E) Mean expression of PI4K2B in different cell types from CRC single-cell samples

diagrams based on patient sex, age, TNM stage, and survival status (Figure 2D). Kaplan-Meier survival curves indicated that the high PI4K2B expression group had significantly better overall survival compared to the low expression group ($p < 0.01$) (Figure 2E).

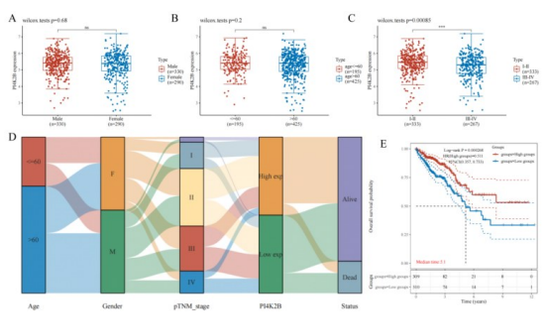

Figure 2 (A) Expression of PI4K2B in CRC tumor samples from patients of different sexes. (B) Expression of PI4K2B in CRC tumor samples from patients of different age groups. (C) Expression of PI4K2B in CRC tumor samples from different tumor stages. (D) Sankey diagram depicting the correlation between PI4K2B and various clinical parameters in CRC. (E) Kaplan-Meier survival curves for high expression group and low expression group of PI4K2B

3.3 PI4K2B as an independent prognostic factor for CRC

We performed univariate (Figure 3A) and multivariate (Figure 3B) Cox regression analyses on PI4K2B and CRC prognosis, generating forest plots that identified PI4K2B as an independent prognostic factor for CRC. Subsequently, nomograms and their calibration curves were plotted (Figure 3C, D), illustrating the contribution of each factor to the multivariate Cox regression model. These nomograms enabled prediction of 1-year, 3-year, and 5-year overall survival rates for individual patients. The model exhibited good predictive accuracy and stability, with a C-index of 0.745.

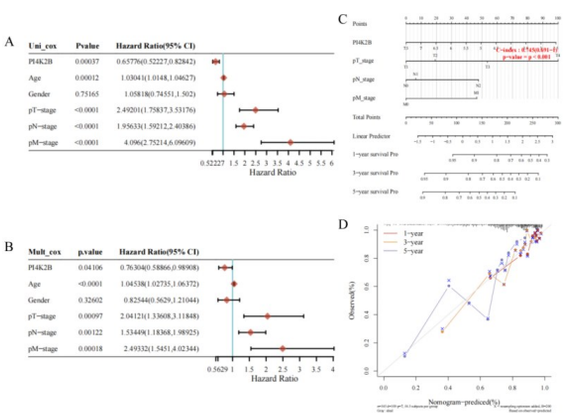

Figure 3 (A) Forest plot of univariate Cox regression analysis for PI4K2B expression, age, sex, and tumor TNM stage in CRC. (B) Forest plot of multivariate Cox regression analysis for PI4K2B expression, age, sex, and tumor TNM stage in CRC. (C) Nomogram for predicting survival probability based on PI4K2B expression, age, sex, and tumor TNM stage in CRC. (D) Calibration curve of the nomogram

3.4 PI4K2B protein interaction network, GSEA, and gene correlation diagram

To further explore the role of PI4K2B in CRC, we constructed a PI4K2B PPI network (Figure 4A). GSEA for PI4K2B revealed its enrichment in 12 pathways: E2F targets, G2M checkpoint, protein secretion, MYC targets V1, mTORC1 signaling, androgen response, unfolded protein response, mitotic spindle, oxidative phosphorylation, spermatogenesis, heme metabolism, and fatty acid metabolism (Figure 4B). We also identified genes in CRC whose expression positively and negatively correlated with PI4K2B expression and generated a correlation network (Figure 4C, D).

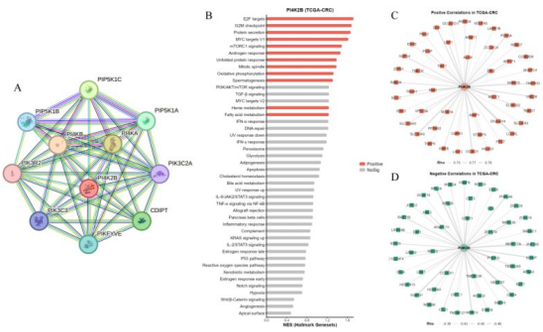

Figure 4 (A) Protein-protein interaction network of PI4K2B. (B) Hallmark gene set enrichment analysis for PI4K2B in CRC. (C) Genes positively correlated with PI4K2B expression in CRC. (D) Genes negatively correlated with PI4K2B expression in CRC

3.5 Differential gene expression analysis of high and low PI4K2B expression groups

We selected 155 samples each from the top 25% (G1) and bottom 25% (G2) of PI4K2B expression levels from 620 CRC samples, forming two groups for differential analysis. In G1, a total of 682 upregulated and 16 downregulated genes were identified (Figure 5A, B). Enrichment analysis revealed that the 682 upregulated genes were mainly enriched in pathways related to cell cycle, nuclear division, and mitosis. Univariate Cox regression analysis was conducted on 698 DEGs, identifying 84 genes associated with CRC prognosis. Risk scores were calculated based on these genes, categorizing 620 CRC samples into high- and low-risk groups. Kaplan-Meier survival curves demonstrated significantly worse overall survival in the high-risk group compared to the low-risk group. Furthermore, a multivariate Cox regression model was constructed using these 84 genes, and an ROC curve was plotted. The model displayed excellent predictive performance, with area under the curve (AUC) values of 0.703, 0.768, and 0.855 for 1-year, 3-year, and 5-year survival, respectively.

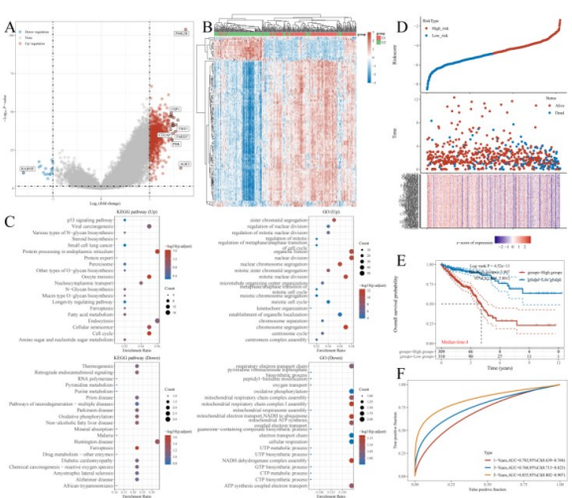

Figure 5 (A) Volcano plot of differentially expressed genes between high PI4K2B expression group (G1) and low expression group (G2). (B) Heatmap of differentially expressed genes in G1 and G2. (C) KEGG and GO enrichment analysis of differentially expressed genes between G1 and G2. (D) Prognostic risk score, survival scatter plot, and gene expression heatmap based on the expression of prognosis-related DEGs between G1 and G2. (E) Kaplan-Meier survival curves for high-risk and low-risk groups. (F) ROC curve for predicting CRC prognosis based on prognosis-related DEGs between G1 and G2

3.6 Relationship between PI4K2B and CRC tumor immune features

In G1, the abundance of neutrophils, macrophages, dendritic cells, B cells, and CD8+ T cells was significantly higher compared to G2 ($p < 0.01$) (Figure 6A). PI4K2B expression exhibited a significant positive correlation ($p < 0.01$) with the abundance of these five cell types (Figure 6B), with the most pronounced association observed with CD8+ T cell infiltration (Figure 6C). TIP analysis (Figure 6D) revealed a non-significant negative correlation between PI4K2B and tumor

immune cell infiltration. Immune checkpoint-related genes CD274, CTLA4, HAVCR2, PDCD1LG2, and TIGIT were significantly upregulated in G1 ($p < 0.05$) (Figure 6E). TIDE algorithm predictions indicated a higher immunotherapeutic response rate in the G1 group, with significantly lower TIDE scores compared to G2 ($p < 0.05$) (Figure 6F).

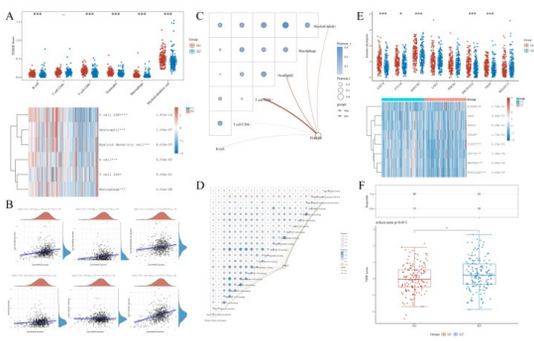

Figure 6 (A) Immune cell infiltration characteristics of G1 and G2. (B) Scatter plot of the correlation between PI4K2B expression and immune cell infiltration. (C) Correlation matrix for PI4K2B expression, immune cell infiltration, and pairwise correlations between immune cells. (D) Correlation matrix for PI4K2B expression, Tumor Immune Phenotype (TIP), and pairwise correlations between TIP components. (E) Expression of immune checkpoints in G1 and G2. (F) TIDE algorithm prediction of the response to immunotherapy in G1 and G2 groups

3.7 Relationship between PI4K2B and CRC drug sensitivity in CRC

Finally, we analyzed the relationship between PI4K2B expression and IC50 values for three CRC therapeutic drugs (5-fluorouracil, irinotecan, and cetuximab), as well as predicting drug treatment responses for G1 and G2. The results indicated a significant negative correlation between PI4K2B expression and IC50 values for all three drugs ($p < 0.01$) (Figure 7). IC50 values for the three drugs were significantly lower in the high PI4K2B expression group compared to the low expression group ($p < 0.01$) (Figure 7), suggesting improved treatment efficacy.

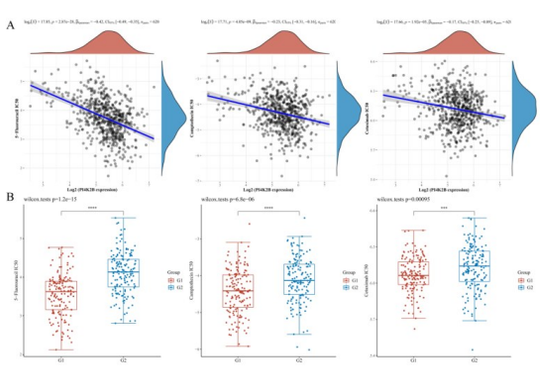

Figure 7 (A) Scatter plot of the correlation between PI4K2B expression and the half-maximal inhibitory concentration (IC50) of 5-fluorouracil, irinotecan, and cetuximab. (B) IC50 values of 5-fluorouracil, irinotecan, and cetuximab in G1 and G2 groups

4 Discussion

The PI4K family has been reported to be closely related to tumorigenesis, development, and treatment[8, 9]. However, there are few reports on the relationship between its important member, PI4K2B, and CRC.

First, we focused on exploring the potential mechanisms underlying PI4K2B's protective role in CRC prognosis. PPIs analysis indicated that PI4K2B interacts with other PI4K family members, PIP5K family member 1, PIP5K1A/C, and PIK3C3. These proteins have been shown to promote cancer cell growth=, invasion, and migration, and they could serve as potential therapeutic targets for cancer=. PI4K2B may exert inhibitory or synergistic effects with these proteins.

GSEA revealed enrichment in pathways such as E2F targets=, G2M checkpoint, and protein secretion, which are associated with cancer cell proliferation, apoptosis, tumor growth, and metastasis. Differential gene expression analysis indicated that DEGs based on PI4K2B expression levels were enriched in pathways related to cell cycle, nuclear division, and mitosis.

We found that PI4K2B expression levels were significantly positively correlated with the abundance of neutrophils, macrophages, dendritic cells, B cells, and CD8+ T cells, and significantly negatively correlated with CD8+ T cell infiltration. This might imply that PI4K2B is involved in regulating immune cell infiltration, particularly playing a crucial role in CD8+ T cell activation or migration. However, TIP analysis showed a non-significant correlation between PI4K2B and tumor immune cell infiltration. It means PI4K2B is not the direct driver of the infiltration of CD8+ T cell but a marker. TIDE

algorithm prediction indicated a higher immunotherapeutic response rate in the high PI4K2B expression group. These findings highlight the potential role of PI4K2B in modulating the tumor immune microenvironment, potentially serving as a reliable target for consideration in CRC immunotherapy.

Finally, we analyzed the relationship between PI4K2B and CRC therapeutic drugs. High PI4K2B expression was significantly negatively correlated with IC50 values for 5-fluorouracil, irinotecan, and cetuximab. IC50 is used to measure cellular sensitivity to drugs, implying that high PI4K2B expression may lead to better CRC treatment outcomes, offering new avenues for CRC treatment.

5 Conclusion

In conclusion, we demonstrate that PI4K2B functions as a Janus-faced player in CRC, with its expression elevated in CRC tissues but acting as a protective factor for CRC patient prognosis.

About the Author

Mingxuan Li, female, MSc, Research Direction: Bioinformatics.

References

[1] Baidoun F, Elshiwy K, Elkeraie Y, Merjaneh Z, Khoudari G, Sarmini MT, Gad M, Al-Husseini M, Saad A: Colorectal Cancer Epidemiology: Recent Trends and Impact on Outcomes. Curr Drug Targets 2021, 22:998-1009.

[2] Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB: Colorectal cancer. Lancet 2019, 394:1467-1480.

[3] Boura E, Nencka R: Phosphatidylinositol 4-kinases: Function, structure, and inhibition. Exp Cell Res 2015, 337:136-145.

[4] Burke JE: Structural Basis for Regulation of Phosphoinositide Kinases and Their Involvement in Human Disease. Mol Cell 2018, 71:653-673.

[5] Dornan GL, McPhail JA, Burke JE: Type III phosphatidylinositol 4 kinases: structure, function, regulation, signalling and involvement in disease. Biochem Soc Trans 2016, 44:260-266.

[6] Govindarajan B, Sbrissa D, Pressprich M, Kim S, Vaishampayan U, Cher ML, Chinni S: Adaptor proteins mediate CXCR4 and PI4KA crosstalk in prostate cancer cells and the significance of PI4KA in bone tumor growth. Res Sq 2023.

[7] Waugh MG: Amplification of Chromosome 1q Genes Encoding the Phosphoinositide Signalling Enzymes PI4KB, AKT3, PIP5K1A and PI3KC2B in Breast Cancer. J Cancer 2014, 5:790-796.

[8] Adhikari H, Kattan WE, Kumar S, Zhou P, Hancock JF, Counter CM: Oncogenic KRAS is dependent upon an EFR3A-PI4KA signaling axis for potent tumorigenic activity. Nat Commun 2021, 12:5248.

[9] Jiang X, Huang X, Zheng G, Jia G, Li Z, Ding X, Lei L, Yuan L, Xu S, Gao N: Targeting PI4KA sensitizes refractory leukemia to chemotherapy by modulating the ERK/AMPK/OXPHOS axis. Theranostics 2022, 12:6972-6988.

[10] Peng JM, Lin SH, Yu MC, Hsieh SY: CLIC1 recruits PIP5K1A/C to induce cell-matrix adhesions for tumor metastasis. J Clin Invest 2021, 131.

[11] Yang G, Song W, Postoak JL, Chen J, Martinez J, Zhang J, Wu L, Van Kaer L: Autophagy-related protein PIK3C3/VPS34 controls T cell metabolism and function. Autophagy 2021, 17:1193-1204.

[12] Shi L, Wang K, Zhao M, Yuan X, Huang C: Overexpression of PIP5KL1 suppresses the growth of human cervical cancer cells in vitro and in vivo. Cell Biol Int 2010, 34:309-315.

[13] Chu CA, Wang YW, Chen YL, Chen HW, Chuang JJ, Chang HY, Ho CL, Chang C, Chow NH, Lee CT: The Role of Phosphatidylinositol 3-Kinase Catalytic Subunit Type 3 in the Pathogenesis of Human Cancer. Int J Mol Sci 2021, 22.

[14] Chu CA, Lee CT, Lee JC, Wang YW, Huang CT, Lan SH, Lin PC, Lin BW, Tian YF, Liu HS, Chow NH: MiR-338-5p promotes metastasis of colorectal cancer by inhibition of phosphatidylinositol 3-kinase, catalytic subunit type 3-mediated autophagy pathway. EBioMedicine 2019, 43:270-281.

[15] Faria M, Domingues R, Bugalho MJ, Silva AL, Matos P: Analysis of NIS Plasma Membrane Interactors Discloses Key Regulation by a SRC/RAC1/PAK1/PIP5K/EZRIN Pathway with Potential Implications for Radioiodine Re-Sensitization Therapy in Thyroid Cancer. Cancers (Basel) 2021, 13.

[16] Andrews DM, Cartic S, Cosulich S, Divecha N, Faulder P, Flemington V, Kern O, Kettle JG, MacDonald E, McKelvie J, et al: Identification and optimization of a novel series of selective PIP5K inhibitors. Bioorg Med Chem 2022, 54:116557.

[17] Yin M, Wang Y: The role of PIP5K1A in cancer development and progression. Med Oncol 2022, 39:151.

[18] Kent LN, Leone G: The broken cycle: E2F dysfunction in cancer. Nat Rev Cancer 2019, 19:326-338.

[19] Sun J, Xia F, Zhang S, Zhang B, Guan Y, Hu X, Xue P, Yang S, Zhou Y, Ling D, Li F: A Selective Nano Cell Cycle Checkpoint Inhibitor Overcomes Leukemia Chemoresistance. Small 2023, 19:e2300736.

[20] Pan S, Chen R: Pathological implication of protein post-translational modifications in cancer. Mol Aspects Med 2022, 86:101097.

[21] Reina-Campos M, Scharping NE, Goldrath AW: CD8(+) T cell metabolism in infection and cancer. Nat Rev Immunol 2021, 21:718-738.

[22] Tichet M, Wullschleger S, Chryplewicz A, Fournier N, Marcone R, Kauzlaric A, Homicsko K, Deak LC, Umaña P, Klein C, Hanahan D: Bispecific PD1-IL2v and anti-PD-L1 break tumor immunity resistance by enhancing stem-like tumor-reactive CD8(+) T cells and reprogramming macrophages. Immunity 2023, 56:162-179.e166.

[23] Liu C, Jin Y, Zhang H, Yan J, Guo Y, Bao X, Zhao P: Effects of KMT2D mutation and its exon 39 mutation on the immune microenvironment and drug sensitivity in colorectal adenocarcinoma. Heliyon 2023, 9:e13629.