KMT2C is Potential Prognostic Biomarker and its Immune Regulating Roles in Pan-Cancer: A Comprehensive Analysis

Guangnan Wei
South China University of Technology

Yuchen Zhang
South China University of Technology

Hongkai Zhuang
Shantou University of Medical College

Yingzi Li
Guangdong Academy of Medical Sciences

Chongyang Ren
Guangdong Academy of Medical Sciences

Lingzhu Wen
Guangdong Academy of Medical Sciences

Jiali Lin
Guangdong Academy of Medical Sciences

Danian Dai
Guangdong Academy of Medical Sciences

Bo Chen
Guangdong Academy of Medical Sciences

Ning Liao (✉ syliaoning@scut.edu.cn)
Guangdong Academy of Medical Sciences

Research Article

Keywords: KMT2C, mutation types, prognosis, immune cells infiltration, pan-cancer

Posted Date: December 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1174621/v1

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

Background: A member of histone lysine methyltransferases subfamily, The histone 3 lysine 4 (H3K4) monomethylase KMT2C, has mutations across many cancer types. However, the role of KMT2C in different cancers and its correlation with tumor infiltration and immune therapy indicators remain unknown.

Method: Expression and mutation information of KMT2C has been analyzed through the Genotype-Tissue Expression (GTEx), The Cancer Genome Atlas (TCGA) Cancer Cell Line Encyclopedia (CCLE) and International Cancer Genome Consortium (ICGC) database in our study. Prognostic value of KMT2C was evaluated via univariate survival analysis and expression detection in different cancer cells.

Result: Survival analysis showed that high expression of KMT2C in some cancer type may be a indication of better outcome, while in other cancer like UVM, patient with high expression of KMT2C suffered from early recurrence. Further, we found there is a strongly link between KMT2C expression and immune cells infiltration, mutation indicators through analysising in the Tumor Immune Evaluation Resource (TIMER) database.

Conclusion: The bioinformatics analysis here deliver us a message that KMT2C might be a good molecular biomarker for prognostic and therapeutic evaluation in specific cancer types.

Introduction

There are ammounting evidencethat epigenetic mechanisms play important roles in progression of many cell types, and confirmations that it acts as a undamental part in cancer onset and development. Epigenetic alterations, including chromatin rearrangement, histone modifications, DNA methylation, and non-coding RNAs during the initiation and progression of disease, have been widely uncovered in recent researches [1,2]. Histone lysine methyltransferases (KMTs), working as a type of fundamental component of chromatin-modifier responsible for active transcription-related histone modifications, has the involvment in methylation process by catalyzing the distribution of histone amino acid sites, thus contributing the structure change of chromatin changed and target genes transcription affections, so as to play an essential role in regulating epigenetic changes [3-5].

Based on primary amino acid sequence and substrate specificity, KMTs are classified into six subfamilies [3]. H3K4 monomethylase KMT2C working as one of member of histone lysine methyltransferases subfamily, has involved in epigenetic modification for gene expressions [6], and KMT2C was located on chromosome 7q36, responsible forencoding nuclear protein with an AT hook DNA-binding domain, which is a protein domain like SET domain, and zinc finger. [7-9].KMT2C acting as a tumor suppressor gene was evidenced by several observations and experimental data, where researches were focused on exploring the function of the SET domain in tumors via knockout models [10-13]. For example, The case that a mouse model of pancreatic adenocarcinoma was constructed by knocking out the SET domain of KMT2C, which contributed to epithelial tumors formation ,also proved that KMT2C may function as a
tumor suppressor gene [14]. Our previous study elucidated KMT2C mutations have high-rate occurrences in Chinese patients with breast cancer and with significant KMT2C mutation differences according to race and ethnicity.[13]

Due to the improvement of bioinformatics technology and the establishment of open-access databases, KMT2C mutation was detected in most of cancer types. However, there were lack of data about KMT2C expression in different tissues and no enough date for revealing the correlation between prognosis and KMT2C expression. A link between tumor immune response and epigenetic regulation had been well-known[15-19] And participation of the epigenetic alterations in gene silencing allow tumors to adapt to changes in surrounding microenvironment [19]. The aim of this study was aimed mining the biologicak information by analyzing the expression of KMT2C in different tissue t, and present the relationship between KMT2C expression and immune microenvironment to figured out whether KMT2C mutations are related to prognosis.

**Results**

**The expression level of KMT2C in normal tissue and tumor**

The GTEx database analysis revealed that KMT2C mRNA expression are similar across all tissues in different healthy people (Figure 1A), but with the lowest expression in blood tissue. The expression level of KMT2C in different cancer cell line analysed from data collected in the CCLE database indicated that KMT2C expression level obviously elevated with narrower ranges compared with normal tissue (Figure 1B). Further, we obtained 9 out of 20 cancer types with significantly different expression in the comparison between tumor and peri-tumor from the TCGA database.

Considering that there are fewer normal samples in TCGA, we integrated the data of TCGA with GTEx, and there was also significant difference were found in 21 out of 27 cancers. The bladder urothelial carcinoma (BLCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), uterine corpus endometrial carcinoma (UCEC) and uterine carcinosarcoma (UCS) was found to be with lower expression of KMT2C in cancer cells compared with normal cells (Figure 1C, D). We investigated the mutation frequency and types of KMT2C in different cancer by TCGA. The mutations in those tumors included missense mutations, nonsense mutations, splicing mutations, frame shift insertions or deletions. Some tumors only presented one mutation type, like KICH only detected KMT2C missense mutation (Supplementary Figure 1S). CESC identified a high mutation frequency (19.38%) of KMT2C, followed by UCEC(17.74%), stomach adenocarcinoma(STAD) (12.13%), lung squamous cell carcinoma (LUSC)(10.98%), and adrenal cortical carcinoma (ACC), cholangiocarcinoma (CHOL), mesothelioma (MESO), pheochromocytoma and paraganglioma(PCPG), thymoma (THYM), UCS, uveal melanoma (UVM) were lack of data to analyzed KMT2C mutation. As show in Figure1E, among the patients with CESC, the highest proportion of mutations were missense mutations, followed by nonsense mutations and frame shift mutation, no splicing mutations.
Prognosis value of KMT2C in cancers

To examine the correlation of KMT2C expression levels with overall survival (OS), we performed univariate survival analysis using TCGA database. The hazard ratios for KMT2C were significant for lower grade glioma (LGG), liver hepatocellular carcinoma (LIHC), acute myeloid leukemia (LAML), lung adenocarcinoma (LUAD), kidney renal clear cell carcinoma (KIRC) (Supplementary Figure 2S). And Kaplan-Meier analysis of the survival differences in KIRC and LUAD shown that patients with lower expression of KMT2C had worse outcome (Figure 2A,B).

We conducted another survival analysis for disease-specific survival (DSS) and KMT2C expression among different cancers in consideration of non-cancer related death could affect OS during follow-up. The significant risk effect is KIRC, LUAD and LGG (Supplementary Figure 3S). KIRC patient with lower expression of KMT2C had worse outcome in the following survival analysis, however, LGG patient with lower expression had better outcome (Figure 2C,D). We further performed univariate survival analysis to show correlation between KMT2C expression and disease-free interval (DFI). The result showed that significant hazard ratios were found in BLCA and ovarian serous cystadenocarcinoma (OV) (Supplementary Figure 4S). In OV, patients with lower expression KMT2C suffered early recurrence after surgery (Figure 2G). Finally, Cox regression was used to illustrate the correlation between KMT2C expression and PFI. There are significant hazard ratios occurred in ACC, KIRC, OV, skin cutaneous melanoma (SKCM) and UVM, among which UVM had the highest risk (Supplementary Figure 5S). In KIRC, high expression patient had better outcome, but in UVM, patient with high expression had worse outcome (Figure 2E,F).

Correlation of KMT2C expression with tumor immune microenvironment and immune checkpoint genes across all tumors

In account of heterogeneousness of immune infiltration of different tumors with KMT2C expression levels may affect clinical results. We downloaded the score data of six immune infiltrating cells from 33 cancers from the TIMER database to performe correlation analysis between KMT2C expression and immune infiltration level. The three top-ranking tumor cohorts were colon adenocarcinoma (COAD), KIRC, and LUAD show high level of KMT2C expression, which has linkages with possible increased infiltration level by immune cells (Figure 3A). These three tumors are infiltrated with different types of immune cells, within which macrophages had the highest significant coefficients in COAD, neutrophils showed the highest significant coefficients in KIRC, in LUAD had a high infiltration CD4+T cells had a high infiltration ratio in LUAD.

In addition, we show the correlation of KMT2C expression with immune score, stroma score and ESTIMATE score across 33 tumor types. For the immune score, the three top-ranking tumor cohorts were sarcoma (SARC), thyroid carcinoma (THCA) and breast invasive carcinoma (BRCA), while in stroma score...
that SARC, head and neck squamous cell carcinoma (HNSC) and glioblastoma multiforme (GBM) were most significant relation. The different result showed in ESTIMATE score that the three top-ranking tumor cohorts were SARC, THCA, and GBM (Figure 3B). To further understand the relationship between the expression level of KMT2C and immune checkpoint genes, we collected acknowledged immune checkpoint genes, and calculated correlation of its expression with our target gene. As shown in Figure 4, NRP1, ADORA2A, CD160, TNFSF15 had a high correlation (P<0.05) with KMT2C in most of cancer types. Data also showed co-expression of KMT2C with more immune checkpoint genes was detected in COAD and LICH. A different result found KMT2C expression was negatively correlated with most immune checkpoint genes in CHOL, esophageal carcinoma (ESCA), MESO, UCS.

**KMT2C is associated with the Neoantigen TMB and MSI in some cancers**

Neoantigen is a new protein encoded by a mutated gene, which included gene point mutations, deletion mutations, gene fusions, in a tumor cell rather than those expressed by normal cells[29]. Here we counted the number of neoantigens in each cancer sample and described correlation between KMT2C expression and number of antigens. In BRCA, KIRC, THCA, prostate adenocarcinoma (PRAD) and LGG, the expression of KMT2C was positively correlated with neoantigens (Figure 5A). Some studies have been reported that microsatellite instability (MSI) status and the level of tumor mutation burden (TMB) are considered to be a robust prognostic biomarkers and therapeutic response to immune checkpoint inhibitors [30]. We performed their respective relationships with KMT2C expression in many kinds of cancer to explore the relevance between KMT2C expression and mutation across tumors. The correlation between KMT2C expression and MSI were obtained significance (P<0.05) from 12 cancer types including lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), HNSC, LUAD, LUSC where LUSC had the highest coefficients, while DLBC had the lowest coefficients (Figure 5B). The coefficient values would demonstrate that KMT2C expression positively correlates with MSI for LUSC, but negatively correlates with MSI for DLBC. When it comes to TMB, 6 out of 33 cancer types (BRCA, KIRC, THCA, THYM, UCEC, UVM) showed a relationship between TMB and KMT2C expression. THYM, UCEC had the highest coefficients, suggesting that KMT2C expression positively relates with high mutation status in those cancer types, but low mutation in BRCA, KIRC, THCA, UVM (particularly THCA).

**KMT2C expression may affect methylation after transcription by regulating MMR genes**

We also investigated the relevance between KMT2C expression level and some common known mismatch repair (MMR) genes (MLH1, MSH2, MSH6, PMS2 and EPCAM). The expression of KMT2C is
closely related to the MMR genes (Figure 6A). In some specific cancer types, the expression of KMT2C had positively relationship with MLH1, MSH2, MSH6 and PMS2. With regard to correlation between the KMT2C expression with four methylation transferases (DNMT1, DNMT2, DNMT3A, DNMT3B), it manifested there is an co-expression in many kind of cancer types except USC. SARC, testicular germ cell tumor (TGCT) demonstrated significantly low co-expression coefficients, but in other cancer type coefficients were all high (Figure 6B).

**Discussion**

Our result have revealed that the expression of KMT2C mRNA varied across tissue by analyzing multiple kinds of cancer from GTEx and TCGA database, and it show especially low expression in blood tissue. When comparing cancer cells with corresponding normal tissues, the KMT2C expression level was obviously elevated among tumors. While mutation frequency of KMT2C differ from cancer types to cancer types. Our study detected the most common mutation types are missense mutation and frameshift mutations, but not all tumors found KMT2C mutation. From COSMIC database, 28.3% of KMT2C are primarily frameshift and missense mutations, a substantial proportion of mutations is missense mutations (17.1%) [7]. There have been reported that the majority of the protein usually truncate by these two types mutations of frameshift or missense, and frameshift mutations probably affected carboxy-terminal SET domain, while missense mutations were found in PHD domain. The different mutation patterns imply that the function of KMT2C might be disturbed by differently localized mutations [31,32].

Interestingly, the survival outcome is not always associated with the mutation frequency or mutation types. For example CESC patients with high KMT2C mutation frequency not suffered from early recurrence of tumor. KMT2C is acted as a tumor suppressor and deletion of the KMT2C gene was shown to be associated with worse outcome in KIRC, LUAD, OV, UVM, while LGG patients with lower expression of KMT2C exhibited better outcome. The explanation for this phenomenon is that KMT2C heterozygous mutations might contribute to heritability incidence. The expression of KMT2C might not be an independent risk factor to predict the prognosis of cancer patients and demonstrated different effects or different mutations on KMT2C function. Genome-wide functional studies showed that KMT2C involved in installation of mono- and dimethylation of H3K4 at gene enhancers, which would help to active gene [6]. When promoter bound to KMT2C, gene transcription can be repressed [33]. It seems that KMT2C play a complex role in tumorigenesis and cancer progression. Epigenetic variations can be affected by diet, exercise, or other factors, and it can be reflected on differences of outcomes in diseases caused by DNA damage [1-2]. However, the expression differences of KMT2C in cancer cells and normal cells discovered in our correlation analysis also leading a different outcome in cancer patients, make it promising biomarker for cancer monitoring and further investigation.

There are another question should be considered is whether KMT2C expression related to immune response or not. In our study, we try to answer this question through analyzing a relationship between its expression and immune microenvironment. We found that KMT2C expression was associated with tumor
immune cell infiltration across different cancer types, the highest scoring tumors were COAD, KIRC, and LUAD. In these three cancer types, the highest coefficients of immune cell infiltration are varied. Studies reveal that tumor-infiltrating lymphocyte is an essential component of the immune microenvironment, working as a mediator in antitumor immunity [34,35]. Macrophages provide the first line of defense against tumor immunity by producing tumor-promoting cytokines.

Previous studies showed that macrophages produced proinflammatory cytokines, among which NF-kB acts as a central regulator, leading to tumorigenesis [36]. Neutrophils are the predominant infiltration cell in early inflammation, involving in initiating and expanding inflammatory response [37]. However, there was lack of data to show the role of KMT2C in immune microenvironment. Our study is consistent with previous studies that KMT2C impacted immune cell infiltration. Moreover, we showed the KMT2C expression have a closely connection with immune checkpoint genes across tumors. However, given the complexity of KMT2C function, there are still other undiscovered mechanisms. Previous studies have focused on proving MSI status and high level of TMB working as a predictor for immune checkpoint inhibitor [30]. According to our result, LUSC patient with the high MSI shown better checkpoint inhibitor responses. And TMB is now recognized as an indicator for screening patient who benefit from immune therapy [30]. The higher TMB achieve good response for immune therapy. In UCEC patient, both TMB and MSI have positively relationship with KMT2C expression in our study, so as to lead to our proposition that KMT2C might be a candidate for predicting immune responses more credible. The peptide fragments that enzymatic hydrolysised from neoantigens are presented as antigens to T cells through DC cells, which can promote T cells to become mature activated T cells for tumor neoantigens specifically recognition and proliferation of these activated T cells [29]. We can design and synthesize neoantigen vaccines according to the mutations of tumor cells induced by the immune activity of tumor neoantigens, and then immunize patients to achieve therapeutic effects [38]. In BRCA, KIRC, THCA, PRAD and LGG, patient might have a good response to neoantigens vaccines.

In normal tissues, the MMR systems recognize and correct DNA mismatch to retain genomic stability [39,40]. Patients who were detected with MMR genes mutation or delete may increase the chances of genetic errors, thus contributing to cancer incidence for genome instability. For example, BRCA germline mutation predispose to breast cancer or ovarian cancer.

We also illustrated the expression of KMT2C was closely related with MMR genes in all cancer types. Recent research has shown that DNA methylation is a well-known epigenetic feature of tumor development [3-5]. In the present study, we found a strongly correlation with KMT2C expression and methyltransferases in human cancer except USC. These results demonstrated that KMT2C involved in tumorigenesis and progression through mediating MMR gene mutation and DNA methylation.

Our results have illustrated the role of KMT2C in tumor development, and provided the useful data for immune therapy across cancer types, but some limitation need to be mentioned. Our study is based on data from public databases with no confirmation through experimentation, it may has biases resulting from confounding factors. Besides, in this article we only showed the KMT2C mRNA level that data
downloaded from open-access databases, the levels of functional protein haven't mentioned. It can't be concluded the influence of post-transcription modification in normal or cancer cells.

Conclusions

In summary, KMT2C is expressed in many tissues, and its high expression found in cancer cells. The expression level of KMT2C is related to survival outcome, for example, KIRC with low expression predict better outcome. In some specific cancer types, KMT2C expression strongly relates to tumor infiltration and immune checkpoint genes expression, it could be affected by the immune microenvironment. The connection between KMT2C expression and immune therapy indicators, like TMB, MSI also reflect its important role in immune therapy. An interesting finding is KMT2C expression can impact the MMR gens and DNA methylation. Further studies of KMT2C would help to better understand the role of KMT2C in human cancer.

Materials And Methds

Samples information and KMT2C expression analysis in Human Pan-Cancer

Data of KMT2C expression in different cancer types was obtained from database, including The Cancer Genome Atlas (TCGA) database and Broad Institute Cancer Cell Line Encyclopedia (CCLE) database. For data of 31 normal tissues (liver, lung, kidney, brain, bone marrow, etc.) collected from the Genotype-Tissue Expression (GTEx) program and downloaded through the GTEx portal [20-22], input the rma function within the R package (R studio version: 1.2.1335, R version: 3.6.1), we can normalize all expression data through log2 conversion[23,24]. Moreover, we also extracted mutation information and mismatch repair (MMR) genes in 33 cancer types via TCGA database. Information of TMB and MSI were both downloaded from TCGA database.

Prognosis analysis

Correlation of KMT2C expression and patients’ prognosis, which included overall survival (OS), disease-free interval (DFI), disease-specific survival (DSS) and progression-free interval (PFI) across 33 cancer types of tumors, were analysed and shown by deploying survival ROC and survival in the R package (rdocumentation.org/packages/survival), and the results also manifested the specificity and time-dependent sensitivity of survival [25]. Kaplan–Meier method was used to perform patients’ survival curves in different cancer type after classifying two group of expression of KMT2C. Hazard ratio (HR) with 95% confidence intervals and log-rank P-value were calculated by univariate survival analysis.

Immune Correlation Analysis
TIMER database based on TCGA was used to analyze the abundance of six subgroups of TIICs including B cells, CD4+T cells, CD8+T cells, neutrophils, dendritic cells, macrophages. [26,27]. The correlation between KMT2C expression and the scores of immune cells was analyzed using the method of spearman correlation analysis. In addition, a link between KMT2C expression and immune checkpoint genes was presented by correlation modules.

Statistics

We used Kruskal–Wallis test to analyze KMT2C expression levels in different normal tissues and different cancer cell lines. When comparing differences, Paired t tests or the t test was performed in comparing the differences of KMT2C expression levels in tumor tissues and normal tissues. Univariate survival analysis was applied to analyze the connection between KMT2C expression and patients’ survival. We utilized Kaplan–Meier methods to compare survival with different levels of KMT2C expression. In this study, Pearson correlation analysis was utilized to evaluate the relevance between KMT2C expression and DNA methyltransferases, MMR. We defined significant and positive correlation when P < 0.05 and r > 0.20. All graphs were established by the R packages of ggplot2 and forestplot[28].

Abbreviations
| Abbreviation | Description |
|--------------|-------------|
| ACC          | adrenalcortical carcinoma |
| BLCA         | bladder urothelial carcinoma |
| BRCA         | breast invasive carcinoma |
| CCLE         | Cancer Cell Line Encyclopedia |
| CESC         | Cervical squamous cell carcinoma and endocervical adenocarcinoma |
| CHOL         | cholangiocarcinoma |
| COAD         | colon adenocarcinoma |
| DFI          | disease-free interval |
| DLBC         | lymphoid neoplasm diffuse large B-cell lymphoma |
| DSS          | disease-specific survival |
| DNMT         | DNA methyltransferase |
| ESCA         | esophageal carcinoma |
| GBM          | glioblastoma multiforme |
| GTEx         | The Genotype-Tissue Expression |
| HNSC         | head and neck squamous cell carcinoma |
| H3K4         | The histone 3 lysine 4 |
| ICGC         | International Cancer Genome Consortium |
| KMTs         | Histone lysine methyltransferases |
| KICH         | kidney chromophobe |
| KIRC         | kidney renal clear cell carcinoma |
| KIRP         | kidney renal papillary cell carcinoma |
| LAML         | acute myeloid leukemia |
| LGG          | brain lower grade glioma |
| LIHC         | liver hepatocellular carcinoma |
| LUAD         | lung adenocarcinoma |
| LUSC         | lung squamous cell carcinoma |
| MESO         | mesothelioma |
| MMR          | mismatch repair |
| MSI          | microsatellite instability |
| Abbreviation | Full Form |
|--------------|-----------|
| OS           | overall survival |
| OV           | ovarian serous cystadenocarcinoma |
| PAAD         | pancreatic adenocarcinoma |
| PCPG         | pheochromocytoma and paraganglioma |
| PD-1         | programmed cell death protein 1s |
| PRAD         | prostate adenocarcinoma |
| READ         | rectum adenocarcinoma |
| ROC          | Receive Operating Characteristic |
| SARC         | sarcoma |
| SKCM         | skin cutaneous melanoma |
| STAD         | stomach adenocarcinoma |
| TCGA         | The Cancer Genome Atlas |
| TGCT         | testicular germ cell tumor |
| THCA         | thyroid carcinoma |
| THYM         | thymoma |
| TIMER        | Tumor IMmune Estimation Resource |
| TMB          | tumor mutation burden |
| TNF          | tumor necrosis factor |
| UCEC         | uterine corpus endometrial carcinoma |
| UCS          | uterine carcinosarcoma |
| UVM          | uveal melanoma |

**Declarations**

**Ethics approval and consent to participate**

All the data included in the analysis are from public databases without the need of permissions from local ethical committees.

**Consent for publication**

Not applicable
Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the TCGA (https://portal.gdc.cancer.gov/) and ICGC (https://icgc.org/) repository.

Competing interests

The authors declare that they have no competing interests.

Funding

The Fundamental Research Funds for the Central Universities (y2syD2192230, Bo Chen); This work was supported by funds from the National Natural Science Foundation of China (81902828, Bo Chen); High-level Hospital Construction Project (DFJH201921, Bo Chen); Medical Scientific Research Foundation of Guangdong Province (B2019039, Bo Chen); Natural Science Foundation of Guangdong Province (2016A030313768, Ning Liao; 2018A030313292) and Research Funds from Guangzhou Municipal Science and Technology Project (201707010418, Ning Liao).

Authors' contributions

NL, GW and BC designed the study. BCHZ and DD analyzed the patient data in TCGA. YZ, JL, GW, YL, CR, LW part in the design and data collection process of the study and were responsible for the content of the manuscript. GW, YZ and BC participated in manuscript writing. All authors read and approved the final manuscript.

Acknowledgement

Not applicable

References

1. Jones PA, Baylin SB. The epigenomics of cancer[J]. Cell, 2007, 128(4): 683–692.
2. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis. 2010,31(1):27-36
3. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science. 2001,293(5532):1068-70
4. Vincent M, Mundbjerg K, Skou Pedersen J, Liang G, Jones PA, Ørntoft TF, Dalsgaard Sørensen K, Wiuf C. epiG: statistical inference and profiling of DNA methylation from whole-genome bisulfite sequencing data. Genome Biol. 2017,18(1):38
5. Klutstein M, Nejman D, Greenfield R, Cedar H. DNA Methylation in Cancer and Aging. Cancer Res. 2016, 76(12):3446-50

6. Vallianatos CN, Iwase S. Disrupted intricacy of histone H3K4 methylation in neurodevelopmental disorders. Epigenomics. 2015;7(3):503-519

7. R.C. Rao, Y. Dou, Hijacked in cancer: the KMT2 (MLL) family of methyltransferases, Nat. Rev. Cancer. 15 (6) (2015) 334–346.

8. Wang L, Zhao Z, Ozark PA, et al. Resetting the epigenetic balance of Polycomb and COMPASS function at enhancers for cancer therapy. Nat Med. 2018;24(6):758-769

9. Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. Proc Natl Acad Sci U S A. 2001,98(23):12902-7

10. Li B, Liu HY, Guo SH, Sun P, Gong FM, Jia BQ. A missense mutation (S3660L) in MLL3 gene influences risk of gastric cancer. J BUON. 2014;19(2):394–397.

11. Li WD, Li QR, Xu SN, Wei FJ, Ye ZJ, Cheng JK, et al. Exome sequencing identifies an MLL3 gene germ line mutation in a pedigree of colorectal cancer and acute myeloid leukemia. Blood. 2013;121(8):1478–1479

12. Sasaki MM, Skol AD, Bao R, Rhodes LV, Chambers R, Vokes EE, et al. Integrated genomic analysis suggests MLL3 is a novel candidate susceptibility gene for familial nasopharyngeal carcinoma. Cancer Epidemiol Biomark Prev. 2015;24(8):1222–1228

13. Chen X, Zhang G, Chen B, Wang Y, Guo L, Cao L, Ren C, Wen L, Liao N. Association between histone lysine methyltransferase KMT2C mutation and clinicopathological factors in breast cancer. Biomed Pharmacother. 2019 Aug;116:108997

14. Lee J, Kim DH, Lee S, Yang QH, Lee DK, Lee SK, et al. A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. Proc Natl Acad Sci U S A. 2009;106(21):8513–8518

15. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. m6A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. Mol Cancer. 2020,19(1):53

16. Li J, Yuan S, Norgard RJ, Yan F, Sun YH, Kim IK, Merrell AJ, Sela Y, Jiang Y, Bhanu NV, Garcia BA, Vonderheide RH, Blanco A, Stanger BZ. Epigenetic and Transcriptional Control of the Epidermal Growth Factor Receptor Regulates the Tumor Immune Microenvironment in Pancreatic Cancer. Cancer Discov. 2020,11:1–18

17. Phan AT, Goldrath AW, Glass CK. Metabolic and Epigenetic Coordination of T Cell and Macrophage Immunity. Immunity. 2017,46(5):714-729

18. Li J, Byrne KT, Yan F, Yamazoe T, Chen Z, Baslan T, Richman LP, Lin JH, Sun YH, Rech AJ, Balli D, Hay CA, Sela Y, Merrell AJ, Liudahl SM, Gordon N, Norgard RJ, Yuan S, Yu S, Chao T, Ye S, Eisinger-Mathason TSK, Faryabi RB, Tobias JW, Lowe SW, Coussens LM, Wherry EJ, Vonderheide RH, Stanger...
BZ. Tumor Cell-Intrinsic Factors Underlie Heterogeneity of Immune Cell Infiltration and Response to Immunotherapy. Immunity. 2018,49(1):178-193.e7

19. Topper MJ, Vaz M, Chiappinelli KB, DeStefano Shields CE, Niknafs N, Yen RC, Wenzel A, Hicks J, Ballew M, Stone M, Tran PT, Zahnow CA, Hellmann MD, Anagnostou V, Strissel PL, Strick R, Velculescu VE, Baylin SB. Epigenetic Therapy Ties MYC Depletion to Reversing Immune Evasion and Treating Lung Cancer. Cell. 2017,171(6):1284-1300.e21

20. Hudson, T.J., Anderson, W., Aretz, A., Barker, A.D., Bell, C., Bernabé, R.R. et al. (2010) International network of cancer genome projects. Nature 464, 993–998

21. Tomczak, K., Czerwinska, P. and Wiznerowicz, M. (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp. Oncol. 19, A68–A77

22. Nusinow, D.P., Szpyt, J., Ghandi, M., Rose, C.M., McDonald, E.R., Kalocsay, M. et al. (2020) Quantitative proteomics of the cancer cell line encyclopedia. Cell 180, 387.e316–402.e316

23. R Core Team (2012) R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/

24. RStudio Team (2015) RStudio: integrated development for R, RStudio, Inc., Boston, MA, U.S.A., http://www.rstudio.com/ 19 Heagerty, P.J., Lumley, T. and Pepe, M.S. (2000) Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 56, 337–344, https://doi.org/10.1111/j.0006-341x.2000.00337x

25. Heagerty, P.J., Lumley, T. and Pepe, M.S. (2000) Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 56, 337–344

26. Li, B., Severson, E., Pignon, J.C., Zhao, H., Li, T., Novak, J. et al. (2016) Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol. 17, 174, https://doi.org/10.1186/s13059-016-1028-7

27. Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J.S. et al. (2017) TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res. 77, e108–e110, https://doi.org/10.1158/0008-5472.CAN-17-0307

28. Wickham, H. (2016) Ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York

29. Peng, M., Mo, Y., Wang, Y. et al. Neoantigen vaccine: an emerging tumor immunotherapy. Mol Cancer. 2019,18(1):128-141

30. Overman, M.J., McDermott, R., Leach, J.L., Lonardi, S., Lenz, H.J. et al. (2017) Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. Lancet Oncol. 18, 1182–1191

31. Mann KM, Ward JM, Yew CC, Kovochich A, Dawson DW, Black MA, et al. Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. Proc Natl Acad Sci U S A. 2012;109(16):5934–5941

32. March HN, Rust AG, Wright NA, ten Hoeve J, de Ridder J, Eldridge M, et al. Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. Nat Genet. 2011;43(12):1202–1209
33. Cheng J, Blum R, Bowman C, et al. A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. Mol Cell. 2014;53(6):979–992

34. Woo J. R., Liss M. A., Muldong M. T., Palazzi K., Strasner A., Ammirante M., et al. (2014). Tumor infiltrating b-cells are increased in prostate cancer tissue. J. Transl. Med. 12:30. 10.1186/1479-5876-12-30

35. Berntsson J., Nodin B., Eberhard J., Micke P., Jirström K. (2016). Prognostic impact of tumour-infiltrating b cells and plasma cells in colorectal cancer. Int. J. Cancer. 139:1129–1139

36. Chen Y, Wang SX, Mu R, Luo X, Liu ZS, Liang B, Zhuo HL, Hao XP, Wang Q, Fang DF, Bai ZF, Wang QY, Wang HM, Jin BF, Gong WL, Zhou T, Zhang XM, Xia Q, Li T. Dysregulation of the miR-324-5p-CUEDC2 axis leads to macrophage dysfunction and is associated with colon cancer. Cell Rep. 2014 Jun 26;7(6):1982-93

37. Junttila M. R., de Sauvage F. J. (2013). Influence of tumour micro-environment heterogeneity on therapeutic response. Nature. 501:346–354

38. Li L, Goedegebuure SP, Gillanders WE. Preclinical and clinical development of neoantigen vaccines. Ann Oncol. 2017;28(suppl_12):xii11-xii17

39. Indraccolo S, Lombardi G, Fassan M, Pasqualini L, Giunco S, Marcato R, Gasparini A, Candiotto C, Nalio S, Fiduccia P, Fanelli GN, Pambuku A, Della Puppa A, D’Avella D, Bonaldi L, Gardiman MP, Bertorelle R, De Rossi A, Zagonel V. Genetic, Epigenetic, and Immunologic Profiling of MMR-Deficient Relapsed Glioblastoma. Clin Cancer Res. 2019,25(6):1828-1837

40. Yurgelun MB, Kulke MH, Fuchs CS, Allen BA, Uno H, Hornick JL, Ukaegbu CI, Brais LK, McNamara PG, Mayer RJ, Schrag D, Meyerhardt JA, Ng K, Kidd J, Singh N, Hartman AR, Wenstrup RJ, Syngal S. Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer. J Clin Oncol. 2017,35(10):1086-1095

**Figures**
Figure 1

KMT2C mRNA expression levels in different normal tissue and tumors

A. Different expression levels of KMT2C in normal tissues data from GTEx database. B. KMT2C expression in tumor cell lines data from CCLE database. C. Difference of KMT2C expression in peri-tumor and tumor data from TCGA database. D. KMT2C expression level differed in cancer cells, peri-tumor cells
and normal cells, combining data from TCGA and GTEx databases. E. Diagram of KMT2C mutation location in CESC.

**Figure 2**

The different levels of KMT2C mRNA expression related to different survival outcome in specific tumors. A. KMT2C expression level related to OS in KIRC. B. KMT2C expression level related to OS in LUAD. C. KMT2C expression level related to DSS in KIRC. D. KMT2C expression level related to DSS in LGG. E. KMT2C expression level related to PFI in KIRC. F. KMT2C expression level related to PFI in UVM. G. KMT2C expression level related to DFI in OV.

**Figure 3**

The highest relevance of KMT2C mRNA expression level and infiltration scores of six common-type immune cells (B cell, CD4+ T cell, CD8+ T cell, neutrophil, macrophage, dendritic cell), StromalScore, Est_ImmuneScore, ESTIMAEScore, counted by TIMER from TCGA database. A. KMT2C mRNA expression level relate to infiltration scores of six immune cells in COAD, KIRC, LUAD. B. The relationship between KMT2C mRNA expression level and StromalScore, Est_ImmuneScore, ESTIMAEScore.
Figure 4

mRNA expression of KMT2C co-expression with immune checkpoint genes in many cancer types from TCGA database. The upper triangle reference log10 transformed P-value and the lower triangle in every tile indicates coefficients calculated by Pearson’s correlation test. *P<0.05, **P<0.01, ***P<0.001.
Figure 5

Correlation between KMT2C mRNA expression levels and neoantigen count, TMB, MSI from TCGA database. A. KMT2C mRNA expression levels correlated with neoantigen count. B. Correlation between MSI and KMT2C expression. C. The relationship of TMB and KMT2C expression. Spearman correlation test, P<0.05 was considered significant.
**Figure 6**

**KMT2C mRNA expression level closely related to MMR genes and methylation.** **A.** Relationship between KMT2C mRNA expression and MMR genes. The upper triangle indicates log10 transformed P-value, and the lower triangle in every tile indicates coefficients calculated by Pearson’s correlation test. *: P<0.05, **: P<0.01, ***: P<0.001. **B.** the relevance of KMT2C and methyltransferases (DNMT1 is red, DNMT2 is blue, DNMT3A is green, DNMT3B is purple).
Supplementary Files

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

- SupplementaryFigures1.jpg
- SupplementaryFigures2.jpg
- SupplementaryFigures3.jpg
- SupplementaryFigures4.jpg
- SupplementaryFigures5.jpg