

Research Paper

KMT2A Mutations and High Prevalence of dMMR-associated Mutational Signatures as Prognostic Indicators in Metastatic Colorectal Cancer

Zhihang Han^{1*}, Chuanjun Song^{1,2*}, Dongqing Li^{1*}, Weiyou Zhu¹, Jiukang Sun¹, Jialing Yao¹, Wen Yuan Gan¹, Fufeng Wang⁴, Xiaodong Yang¹✉, Lingjun Zhu^{1,3}✉

1. The First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Nanjing, Jiangsu Province, 210029, China.
2. Xinghua People's Hospital Affiliated to Yangzhou University, No. 419 Yingwu South Road, Xinghua, Jiangsu, 225700, China.
3. Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, Jiangsu Province, 210029, China.
4. Nanjing Geneseeq Technology Inc, Nanjing, Jiangsu Province, 210018, China.

* These authors contributed equally.

✉ Corresponding authors: Xiaodong Yang, Department of General Surgery, The First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Nanjing, 210029, China; Email: 20113276@qq.com. Lingjun Zhu, Department of Oncology, The First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Nanjing, 210029, China; Email: zhulingjun@njmu.edu.cn.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2024.01.18; Accepted: 2024.03.17; Published: 2024.04.08

Abstract

The conventional treatment strategies for patients with metastatic colorectal cancer (mCRC) are predominantly guided by the status of RAS and BRAF mutations. Although patients may exhibit analogous pathological characteristics and undergo similar treatment regimens, notable disparities in their prognostic outcomes can be observed. Therefore, tissue and plasma samples from 40 mCRC patients underwent next-generation sequencing targeting 425 cancer-relevant genes. Genomic variations and canonical oncogenic pathways were investigated for their prognostic effects in association with progression-free survival (PFS) of these patients. We found that patients with *BRCA2* and *KMT2A* mutations exhibited worse prognostic outcomes after chemotherapy-based treatment (univariate, $P < 0.01$). Further pathway analysis indicated that alterations in the homologous recombination pathway and in the *KMT2A* signaling network were also significantly associated with shortened PFS (univariate, $P < 0.01$). Additionally, mutation signature analysis showed that patients with higher proportions of defective mismatch repair (dMMR)-related mutational signatures. Had a worse prognosis (univariate, $P = 0.02$). *KMT2A* mutations (hazard ratio [HR], 4.47; 95% confidence interval [CI], 1-19.93; $P = 0.050$) and dMMR signature proportions (HR, 3.57; 95% CI, 1.42-8.96; $P = 0.007$) remained independently associated with PFS after multivariate analysis and the results were further externally validated. These findings may enhance our understanding of this disease and may potentially facilitate the optimization of its treatment approaches.

Keywords: colorectal cancer, prognostic indicators, next-generation sequencing, *KMT2A* mutations, dMMR-associated Mutational Signatures

Introduction

Colorectal cancer (CRC) ranks as the third most prevalent and the second most lethal malignancy worldwide. The primary cause of mortality in CRC patients is metastasis, which significantly diminishes the prognosis of those affected [1]. Despite advancements in treatment modalities, the five-year survival rate for patients diagnosed with metastatic CRC (mCRC) remains dismally low, ranging between 11% and 15% [2]. There is an urgent need for the

identification and validation of robust prognostic biomarkers that could pave the way for early intervention and broaden the spectrum of therapeutic options for mCRC patients.

The prognostic landscape of CRC has traditionally relied on clinicopathological characteristics, with the stage at diagnosis serving as a pivotal indicator of patient outcomes. Other factors influencing prognosis include tumor localization, the

presence of perineural invasion, and the degree of histological differentiation [3]. In recent years, the advent of molecular biology, alongside breakthroughs in immunotherapy and targeted treatments, has revolutionized the treatment landscape for advanced cancers. This paradigm shift has sparked a growing interest in the exploration of molecular biomarkers as tools for prognostication in CRC.

Microsatellite instability (MSI) is a favorable prognostic factor in stage II CRC but may have a marginal negative impact on survival in the metastatic setting [4-6]. The presence of mutations in the RAS and BRAF genes plays a critical role in determining treatment efficacy and thereby influencing prognostic outcomes in mCRC. Evidence from a meta-analysis encompassing five clinical trials illustrates the correlation between mutations in *KRAS* and *BRAF* and reduced progression-free survival (PFS) and overall survival (OS) rates [7]. In addition, the CpG island methylator phenotype (CIMP), characterized by aberrant methylation patterns, is associated with a shorter OS and varying responses to conventional chemotherapy [8, 9]. The co-occurrence of *KRAS*/*BRAF* mutations or MSI in CIMP-positive tumors complicates the assessment of CIMP's independent prognostic value [10].

Despite these insights, the clinical adoption of prognostic biomarkers in CRC remains limited, underscoring the necessity for further investigation. Leveraging the advantage of next-generation sequencing, this study aims to delve deeper into the prognostic markers for CRC, with the objective of uncovering novel, independent prognostic indicators that could inform therapeutic decisions and enhance clinical outcomes for CRC patients.

Material and Methods

Patient and sample inclusion

This study retrospectively collected data from patients diagnosed with late-stage recurrent or metastatic colorectal cancer (mCRC) who received treatment at Jiangsu Provincial People's Hospital between November 2017 and February 2022. The patient selection criteria were as follows.

Inclusion criteria: 1) Diagnosis of late-stage recurrent or metastatic colorectal cancer. 2) Availability of comprehensive clinical information. 3) Detailed records of treatment received, including medication types, treatment duration, and progression-free survival data. (4) Availability of tumor tissue samples suitable for next-generation sequencing (NGS) analysis.

Exclusion criteria: 1) Incomplete clinical or treatment information. 2) Patients who underwent

surgery after treatment.

According to these criteria, a total of 40 CRC patients were included in the study.

Library preparation and sequencing

Hybridization-based target enrichment of 437 cancer-related genes was carried out using the GeneseeqPrime® pan-cancer gene panel with xGen Lockdown Hybridization and Wash Reagents Kit (Integrated DNA Technologies). Captured libraries by Dynabeads M-270 (Life Technologies) were amplified in KAPA HiFi HotStart ReadyMix (KAPA Biosystems) and quantified by qPCR using KAPA Library Quantification Kit (KAPA Biosystems). The final libraries were sequenced on Hiseq4000 platform (Illumina).

Sequencing data and bioinformatics analysis

Trimmomatic was used for FASTQ file quality control. Leading or trailing low-quality bases ($Q < 20$) and N bases were removed [11]. The sequencing data was aligned to the reference Human Genome (hg19) using Burrows-Wheeler Aligner (BWA-mem, v0.7.12) [12], and was then de-duplicated by Sambamba [13]. Base quality recalibration and indel realignment were processed by Genome Analysis Toolkit (GATK 3.4.0) [14]. VarScan2 [15] was employed for calling single-nucleotide variations (SNVs) and insertion/deletions (indels). Genomic fusions were identified by FACTERA [16] with default parameters. Copy-number variations (CNVs) were detected using CNVkit [17] with default parameters. Somatic CNVs were identified comparing paired normal and tumor samples with the cut-off ratio of 0.6 for copy-number loss and 2.0 for copy-number gain. Pathway analyses referred to the STRING database (<https://www.string-db.org/>) for signaling networks.

Mutation signature analysis

"Sigminer" package (<https://cran.r-project.org/web/packages/sigminer/>) was used to extract mutational signatures [18]. The mutational patterns were compared with the COSMIC mutational signatures v3.1 reported by Alexandrov et al. [19]. The contributions of the signatures sharing the same etiological origins were summed. For example, single base substitution (SBS) signatures 6, 14, 15, 20, 21, 26, and 44, all of which are associated with dMMR, were combined to represent the signature of dMMR.

Statistical analysis

Log-rank tests were used to analyze PFS differences among groups. Multivariate analyses using the Cox proportional hazards model were performed to investigate the association between

patients' survival and their clinical or genomic characteristics. A two-sided *P*-value of less than 0.05 was considered significant. All statistical analyses were performed in R (version 3.6.3).

Results

Patient Characteristics and Clinical Factors on Prognosis

A total of 40 CRC patients were included in the study (Figure S1, Table 1), with a median age of 59 (ranging from 44 to 82 years old). Among these patients, 21 were males (52.5%) and 19 were females (47.5%). The primary tumor sites were distributed as follows: left colon (n=12, 30%), right colon (n=12, 30%), rectum (n=15, 37.5%), and left colon plus rectum (n=1, 2.5%). 30 (75%) patients had surgery history. All patients were classified as stage IV according to the American Joint Committee on Cancer (AJCC) criteria and received chemotherapy-based regions. The majority of patients (21, 52.5%) had received chemotherapy combined with bevacizumab, followed by 13 (32.5%) receiving chemotherapy alone, 5 (12.5%) receiving chemotherapy combined with cetuximab, and 1 (2.5%) receiving chemotherapy combined with anlotinib. The median PFS (mPFS) was 9.4 months and the median follow-up time was 13.7 months. Of all clinical factors, only administration of maintenance therapy was significantly associated with survival (Figure 1A and B) and was thus included in subsequent multivariate analyses.

Table 1. Clinical characteristics of the enrolled CRC patients who had baseline tissue samples (N=40).

| Clinical info | All (N=40) |
|----------------------------------|------------|
| Age (years) | |
| Median(range) | 59(44-82) |
| Sex | |
| Female | 19 (47.5%) |
| Male | 21 (52.5%) |
| Stage | |
| IV | 40 (100%) |
| Location of primary tumor | |
| Left colon | 12 (30.0%) |
| Right colon | 12 (30.0%) |
| Left colon+Rectum | 1 (2.5%) |
| Rectum | 15 (37.5%) |
| Surgical history | |
| Yes | 30 (75.0%) |
| No | 10 (25.0%) |
| Prior systemic therapy | |
| Neoadjuvant therapy | 5 (12.5%) |
| Adjuvant therapy | 8 (20.0%) |
| Therapeutic regimen | |
| Chemo+Beva | 21 (52.5%) |
| Chemo | 13 (32.5%) |
| Chemo+Cetuxi | 5 (12.5%) |
| Chemo+Anlotinib | 1 (2.5%) |
| Maintenance treatment | |
| Yes | 6 (15.0%) |
| No | 34 (85.0%) |

Beva, bevacizumab; Cetuxi, cetuximab; Chemo, chemotherapy.

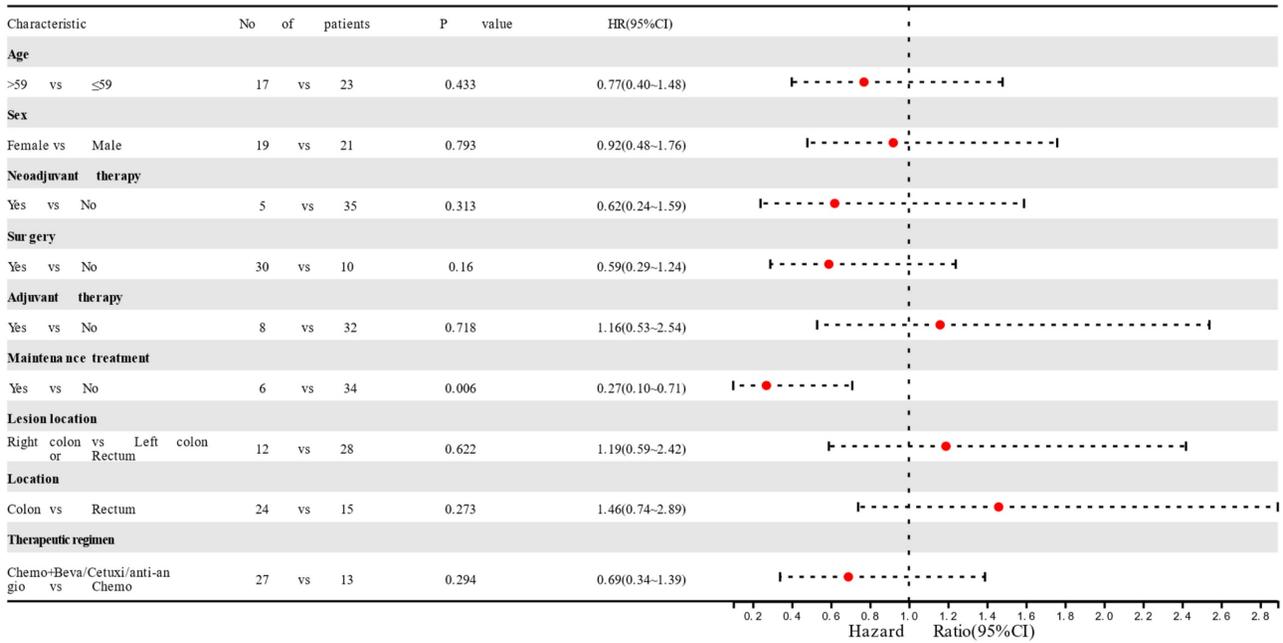
Identification of Prognostic Biomarkers

To identify potential molecular biomarkers associated with outcomes, we profiled the mutational landscape of the studied population (Figure 2A). The top four mutated genes were *TP53* (90%), *APC* (80%), *KRAS* (60%), and *FBXW7* (32.5%), consistent with previous reports on CRC [20]. Only genes that mutated in more than three patients were included in prognosis analysis. Our findings revealed that patients with *BRCA2* or *KMT2A* mutations had a worse prognosis. The mPFS for *BRCA2*-mutated (*BRCA2+*) patients was 4.1 months, compared to 10.0 months for those with wildtype *BRCA2* ($P < 0.001$), showing a hazard ratio (HR) of 7.30 with a 95% confidence level (CI) of 1.88-28.38. Similarly, patients with *KMT2A* mutations had a mPFS of 4.1 months, compared to 10.0 months for those without ($P = 0.012$, HR [95% CI]: 4.41 [1.23-15.82]) (Figure 3A-B). Gene mutations were further grouped based on the signaling pathways or networks in which they are involved, including the top ten cancer-related pathways and the DNA damage response (DDR) pathways [21, 22] (Figure 3F). We found that mutations in the homologous recombination (HR) pathway genes were significantly associated with a poor prognosis, as patients with such mutations had a mPFS of 5.0 months compared to 10.2 months in those without ($P = 0.022$, HR [95% CI]: 2.81 [1.11-7.11]) (Figure 3C, 3F). Significant difference in PFS (5.7 vs 10.0 months) was also observed between patients with and without mutations in genes in the *KMT2A*-related signaling network ($P = 0.021$, HR [95% CI]: 2.68 [1.12-6.39]) (Figure 3D-F).

Moreover, we examined the impact of mutational signatures on the prognosis of 37 patients. The mutational signatures in the cohort were predominantly influenced by age, dMMR, and activation-induced cytidine deaminase (APOBEC) (Figure 4A). We determined an optimized cut-off value for the proportion of dMMR-associated signatures at 1.17×10^{-5} . Patients with a higher proportion of dMMR signature ($>1.17 \times 10^{-5}$) had a significantly worse prognosis, with a mPFS of 6.6 months compared to 12.1 months among those with a lower proportion ($P = 0.020$, HR [95% CI]: 2.30 [1.12-4.71]) (Figure 4B).

Current treatment regimens for CRC patients mainly involve chemotherapy, either as monotherapy or combined with bevacizumab. Considering the relatively high prevalence of alterations in *KMT2A*-related genes and in the HR pathway, further analyses were performed to examine their association with prognosis in patients undergoing these two treatment options.

A



B

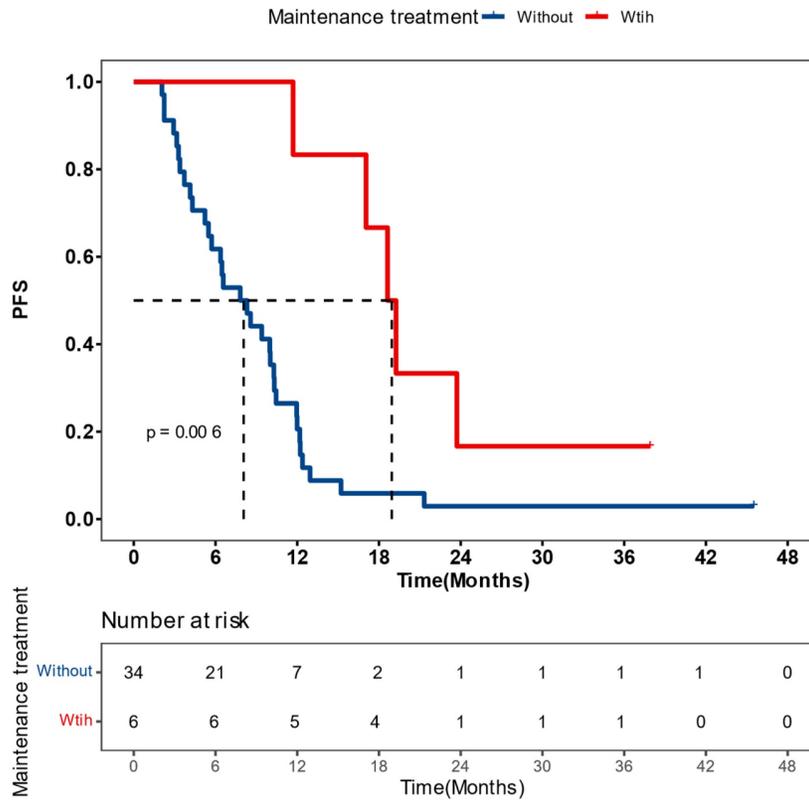


Figure 1. Clinical features on prognosis (A) Cox univariate analysis showing the impact of clinical features on patient progression-free survival (PFS). (B) Kaplan-Meier estimates comparing PFS of patients with or without maintenance therapy. Patients receiving maintenance therapy had significantly longer PFS ($P < 0.05$). Median PFS was indicated by dashed lines. anti-angi, anti-angiogenic therapy; Beva, bevacizumab; Cetuxi, cetuximab; CI, confidence interval; Chemo, chemotherapy; HR, hazard ratio.

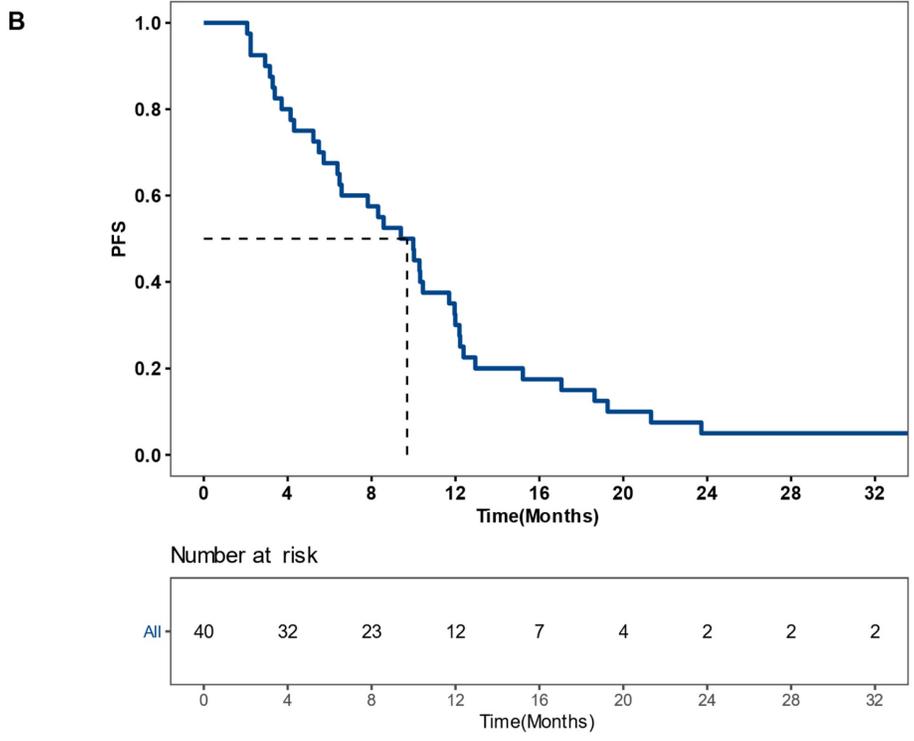
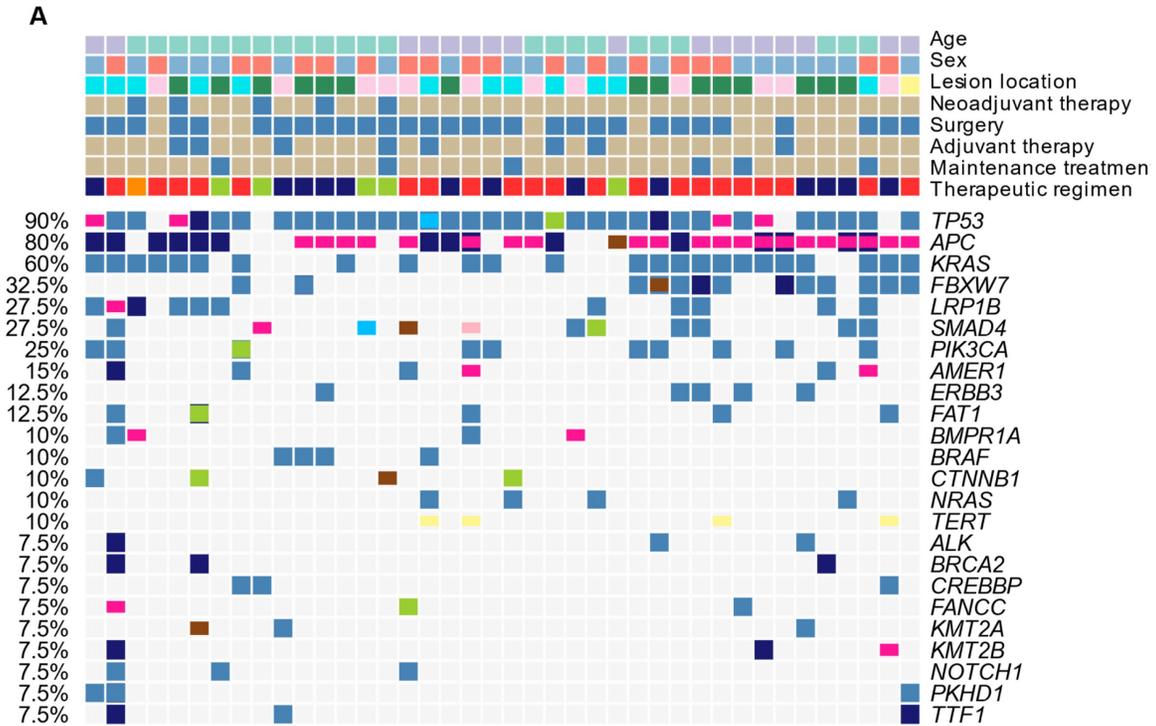


Figure 2. Recurrence risk and mutation profiling of patients (A) The mutational landscape of enrolled patients as detected from their baseline tissue samples. Genes with $\geq 5\%$ mutation prevalence were displayed. (B) A Kaplan-Meier curve showing progression-free survival (PFS) of the studied population. Median PFS was indicated by dashed lines. anti-angio, anti-angiogenic therapy; beva, bevacizumab; cetuxi, cetuximab; Chemo, chemotherapy; SV, structural variant as gene fusion or rearrangement.

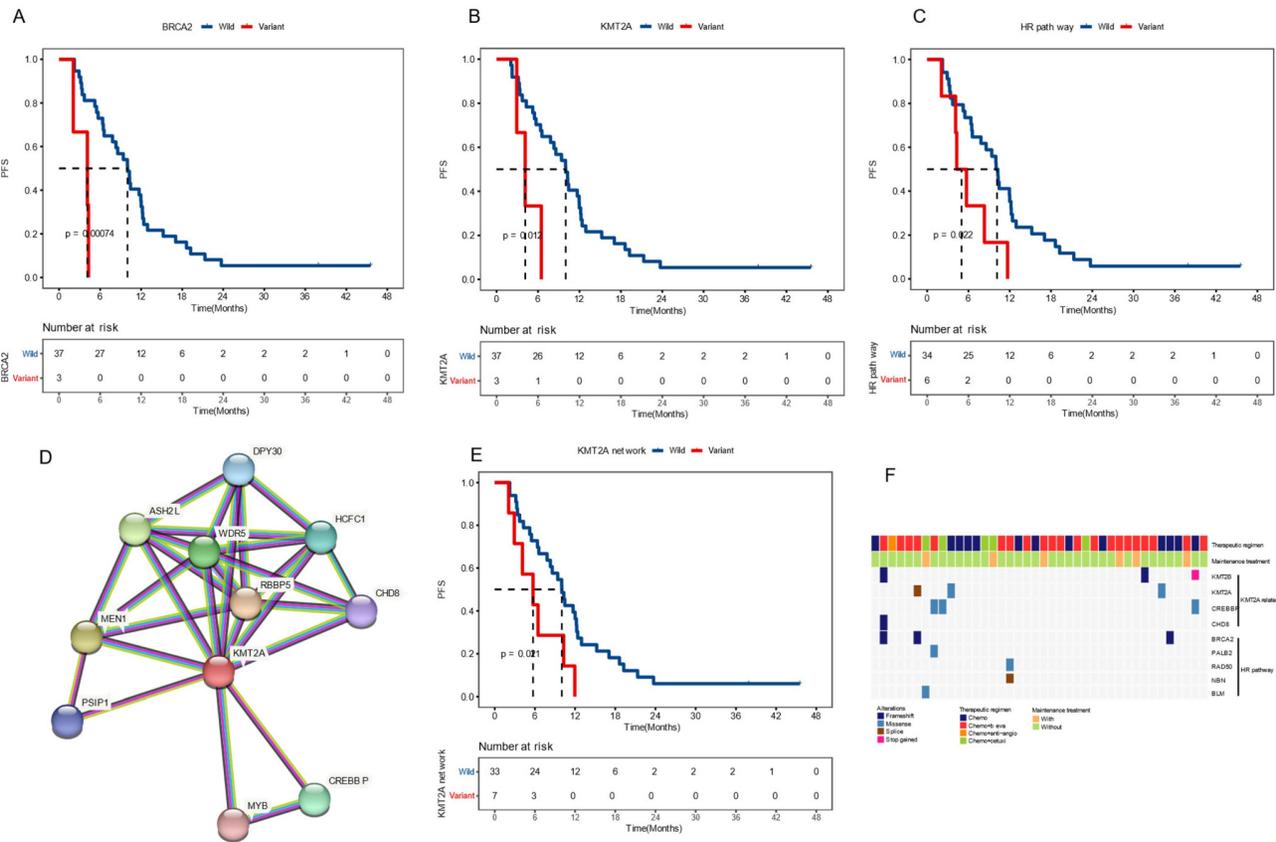


Figure 3. Baseline molecular features on prognosis (A-C) Kaplan-Meier plots of progression-free survival in patients with and without *BRCA2* mutations (A), *KMT2A* mutations (B), and altered homologous recombination (HR) pathway (C). (D) The signaling network associated with *KMT2A* as shown in the STRING database. (E) Kaplan-Meier plots of progression-free survival in patients with and without mutations in the *KMT2A*-related network. (F) The mutational profile of genes in the *KMT2A* network and the HR pathway. Median PFS was indicated by dashed lines. anti-angi, anti-angiogenic therapy; beva, bevacizumab; cetuxi, cetuximab; Chemo, chemotherapy.

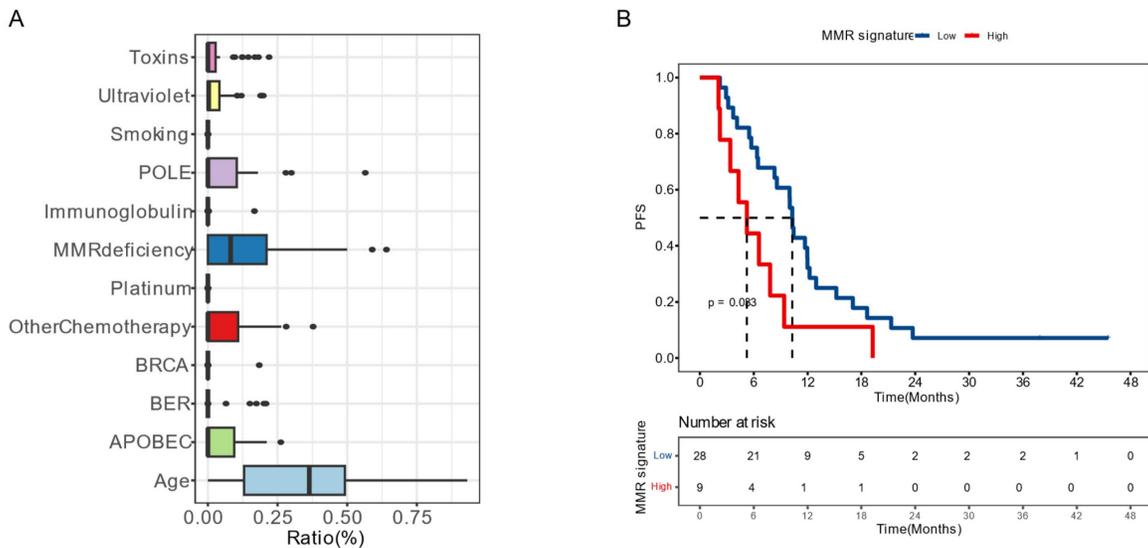


Figure 4. Mutation signatures on patient prognosis (A) Distribution of the mutation signatures clustered from COSMIC (v3.3) single base substitutions. (B) Kaplan-Meier plots of progression-free survival in patients with high or low proportion of deficient mismatch repair (dMMR) signature. Median PFS was indicated by dashed lines. BER, base excision repair.

In the group receiving chemotherapy plus bevacizumab, patients who harbored *KMT2A*-related mutations had a worse prognosis compared to those with the wild-type genes (mPFS 4.1 vs. 9.3 months, $P = 0.02$, HR [95% CI]: 4.48 [1.11-18.13]), and no

significant difference was observed in the group receiving chemotherapy alone (mPFS 6.5 vs. 10.0 months, $P = 0.24$, HR [95% CI]: 2.27 [0.56-9.14]) (Figure 5A). Similarly, patients with HR pathway mutations in the combined therapy group had a

shorter PFS (mPFS 4.9 vs. 10.9 months, $P = 0.029$, HR [95% CI]: 3.01 [0.9-10.11]), while no significant difference was found in the chemotherapy group (mPFS 4.3 vs. 10.0 months, $P = 0.25$, HR [95% CI]: 3.47[0.36-33.44]) (Figure 5B). We also analyzed the impact of dMMR-related signatures on patient prognosis in these two groups, where we found that patients with a higher proportion of dMMR signatures had a worse prognosis in both groups (Figure 5C). We used Fisher's exact test to investigate the interaction of above-mentioned prognosis-related factors, and found significant co-occurrences between mutations in *KMT2A* and those in genes that belong to its signaling network, as well as between *BRCA2* mutations and the HR pathway mutations (Figure S2). Subsequently we performed a multivariate Cox analysis incorporating relevant clinical and genomic features. Considering the co-occurrences, we selected the *KMT2A* and *BRCA2* mutations, which showed stronger correlation with prognosis in univariate analysis, to represent other mutations in their respective pathways. The multivariate analysis identified three independent prognostic factors associated with PFS: maintenance treatment (HR, 0.22[95% CI, 0.08-0.62]; $P = 0.004$), *KMT2A* mutations (HR, 4.47[95% CI, 1-19.93]; $P = 0.050$), and dMMR-related mutational signatures (HR, 3.57[95% CI, 1.42-8.96]; $P = 0.007$) (Table 2). We further examined the association between molecular mutations such as *KMT2A* and clinical variables and found no significant association (Table S1).

External Validation on *KMT2A* Mutation and dMMR Signature

To validate the prognostic effects of *KMT2A* and dMMR signature in advanced CRC, we gathered data from a 2022 study on pan-cancer metastasis

mechanisms by the Memorial Sloan Kettering Cancer Center (MSKCC). The dataset included the mutational and survival information of 2,342 CRC patients who developed distant metastasis and exhibited microsatellite stability (MSS) [23]. We demonstrated that, in this dataset, patients with *KMT2A* mutations had significantly shorter overall survival (OS), with a median OS (mOS) of 26.0 months versus 38.3 months for those without ($P = 0.021$, HR [95% CI]: 1.81 [1.08-3.01]) (Figure 6A). Meanwhile, patients with a higher proportion of dMMR mutational signature had marginally worse OS, with a mOS of 34.3 months compared to 39.0 months in patients with a lower proportion ($P = 0.056$, HR [95% CI]: 1.33 [0.99-1.78]) (Figure 6B). These findings further supported the significance of *KMT2A* mutation and dMMR signature in predicting the prognosis of CRC patients with distant metastasis and MSS.

Discussion

Our study on advanced CRC patients identified several factors potentially predicting a poor prognosis. These include absence of maintenance therapy, genomic variations in *KMT2A* and *BRCA2* and their respective pathways, and enrichment of dMMR-associated mutational signatures. The multivariate Cox analysis revealed that maintenance therapy, *KMT2A* mutations and dMMR mutational signatures remained significantly associated with prognosis, suggesting their potential as independent prognostic indicators. External validation in advanced CRC patients from the MSKCC dataset further supported the negative correlation of *KMT2A* mutations or dMMR signature enrichment with prognosis [23].

Table 2. Cox multivariate analysis of clinical and genomic features affecting patients' progression-free survival.

| | Univariate analysis | | Multivariate analysis | |
|-----------------------|---------------------|---------|-----------------------|---------|
| | HR(95%CI) | P value | HR(95%CI) | P value |
| Maintenance treatment | | | | |
| With vs Without | 0.27(0.10-0.71) | 0.006 | 0.22(0.08-0.62) | 0.004 |
| <i>KMT2A</i> | | | | |
| Variant vs Wild | 4.41(1.23-15.82) | 0.012 | 4.47(1-19.93) | 0.050 |
| <i>BRCA2</i> | | | | |
| Variant vs Wild | 7.30(1.88-28.38) | <0.001 | 2.86(0.63-12.87) | 0.172 |
| MMR signature | | | | |
| Variant vs Wild | 2.30(1.05-5.03) | 0.033 | 3.57(1.42-8.96) | 0.007 |

CI, confidence interval; HR, hazard ratio.

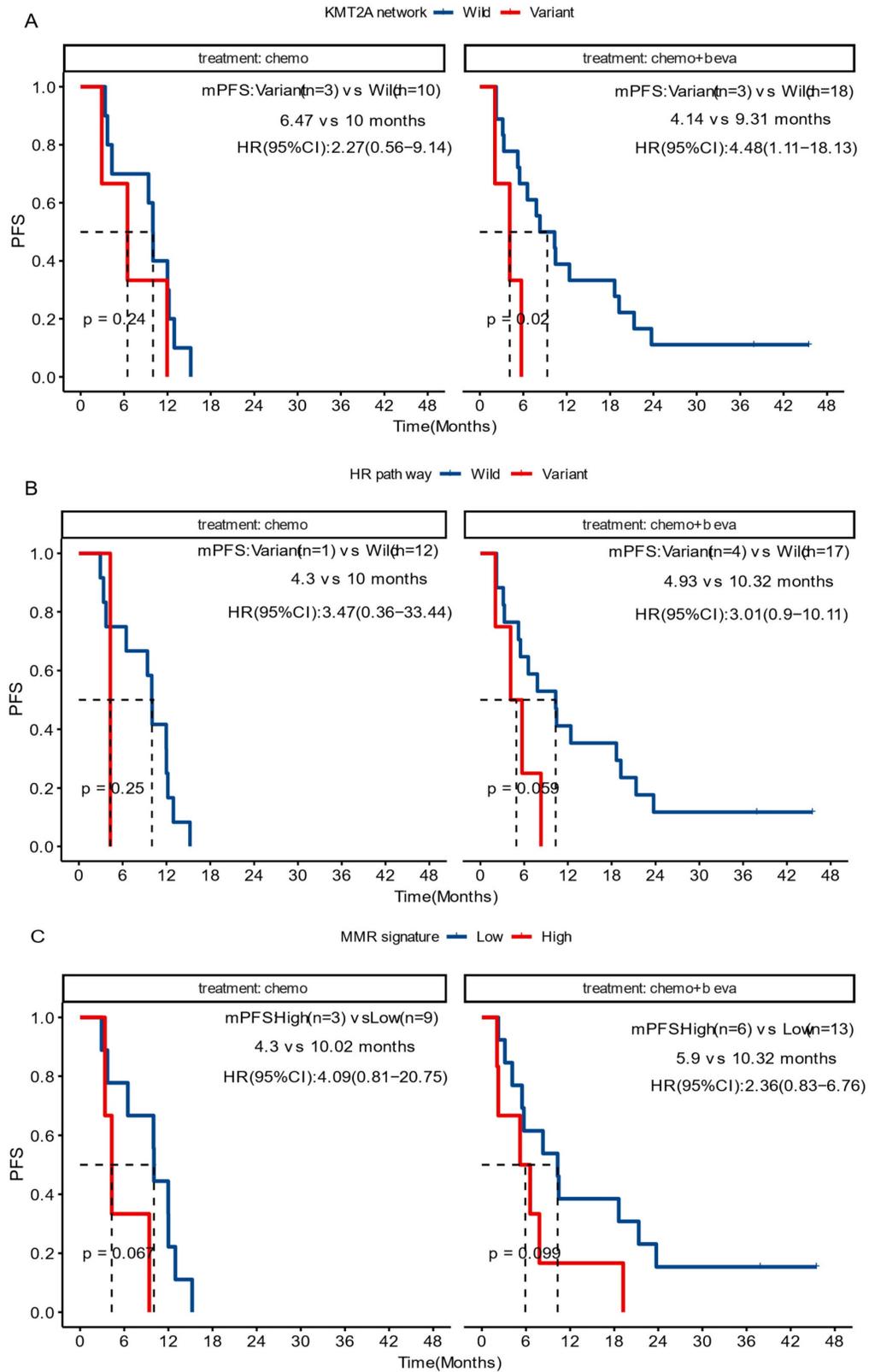


Figure 5. Prognosis among diverse treatment cohorts. The impact of variations in the KMT2A network (A) and in the HR pathway (B) genes, and the impact of dMMR mutation signature (C), on the progression-free survival of patients receiving chemotherapy (chemo) or chemotherapy plus bevacizumab (chemo+beva).

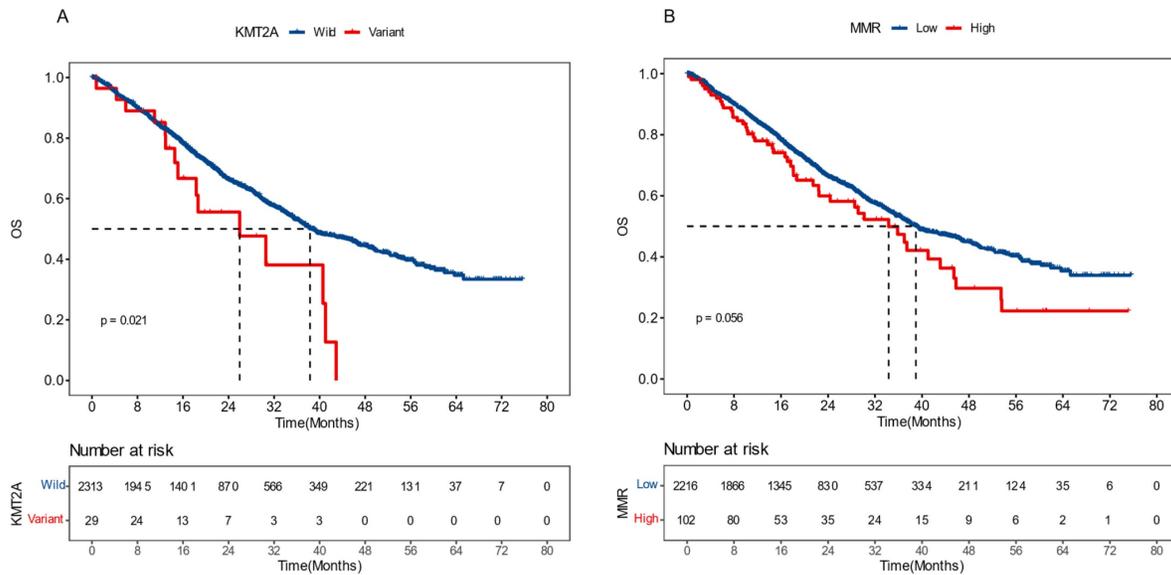


Figure 6. External validation results of KMT2A gene and dMMR signature. The influence of KMT2A mutation (A) and proportion of dMMR signature (B) on patients' overall survival, derived from the Memorial Sloan Kettering Cancer Center dataset [23].

The *KMT2* gene family plays a crucial role in epigenetic regulation, with *KMT2A* specifically first found to be involved in the rearrangement of chromosome 11q23 in mixed-lineage leukemia (MLL) [24, 25]. Recent genomic sequencing studies have shown that *KMT2* genes are commonly mutated in various human cancers [26]. Further studies have revealed that overexpression of *KMT2A* is associated with an unfavorable prognosis in CRC. *KMT2A* mediates the interaction of β -catenin with consensus DNA sequences and the subsequent transcription of β -catenin targets, which consequently promotes tumor growth [27]. A study conducted by Yang Fang and colleagues revealed an overexpression of *KMT2A* in CRC tissues compared with paired paracancerous tissues, with the expression level positively correlated with tumor staging. Moreover, *KMT2A*-knockout cells exhibited suppressed cell migration and invasion [28]. In contrast, Cun Liao *et al.* discovered that *KMT2* variation was associated with increased tumor mutational burden (TMB) and MSI, ultimately resulting in improved survival for CRC patients [29]. In comparison to prior studies, we scrutinized the *KMT2A* protein network via the STRING database and demonstrated that mCRC patients with alterations in *KMT2A*-related genes also exhibited shorter PFS. We further specified that significant PFS difference was only seen in patients receiving chemotherapy plus bevacizumab, but not in those receiving chemotherapy alone. These findings emphasize the need for a comprehensive understanding of the *KMT2A* network and its role in tumor progression and drug response, especially to bevacizumab, a frequently used therapeutic agent in

CRC. Further studies remain to explore the potential of targeting *KMT2A* and its interacting partners as a novel therapeutic strategy for CRC.

Extensive research has been conducted on the prevalence and impact of *BRCA* mutations in cancers. Initial studies have shown that individuals with *BRCA1* or *BRCA2* mutations were 84% more susceptible to breast cancer and 40% more susceptible to ovarian cancer throughout their lifetime [30-32]. *BRCA1/BRCA2* mutations have also been linked to increased risks of prostate, pancreatic, and gastric cancers [33, 34]. A meta-analysis has suggested that individuals with *BRCA1* mutations have a higher risk of developing CRC, but the risk is not higher for those with *BRCA2* mutations; However, the findings of this study are subject to debate [35]. Studies have also indicated that CRC patients with *BRCA* mutations are more likely to benefit from chemotherapy. However, this conclusion was primarily based on case reports and limited retrospective case-control analyses that mainly concerned chemotherapy regimens involving oxaliplatin [36, 37]. In our study, we revealed that mCRC patients with *BRCA2* mutations had a worse prognosis than those without. Treatment outcomes were also suboptimal for patients harboring HR pathway mutations, especially in the cohort receiving chemotherapy plus bevacizumab. HR has been identified as a crucial DNA repair mechanism in mammalian cells, and the indispensable role of the tumor suppressors *BRCA1* and *BRCA2* in HR has been well established [38, 39]. HR deficiency is now considered targetable in cancer treatment as it renders tumor cells more susceptible to certain DNA-damaging agents [40, 41].

The development of CRC is influenced not only by single mutations but also by the combined effect of multiple factors, which yield various genomic patterns termed mutational signatures [42]. An algorithm has been developed to quantify the differential contributions of these mutation signatures to genomic alteration, by assigning them each an activity level, or “exposure” [43]. This provides a more precise understanding of the tumor's genomic feature and the processes underlying its development. Various types of DNA repair mechanisms are activated in response to either endogenous or exogenous DNA damage [44]. The MMR mechanism, in particular, is responsible for correcting base pair mismatches that occur during DNA replication [45] and serves as a physiological barrier against genomic instability and the accumulation of somatic mutations [46]. However, genetic and epigenetic alterations can disrupt these DNA repair pathways, which can be exploited by tumor cells. Consequently, mutations accumulate and their pattern can be recognized by mutational signature analysis [47]. For example, MSI is almost always a direct result of MMR malfunction [48]. The close association between MMR pathway deficiency and both sporadic and hereditary CRC cases, including Lynch syndrome, is well-established [45, 49, 50]. In this study, we found that mutational signatures of the studied mCRC population were primarily determined by age, dMMR, and APOBEC. Additionally, we established a threshold for the proportion of dMMR-related signatures and discovered that patients with a higher proportion tended to have a poorer prognosis. This finding has also been validated by external data. Previous studies have indicated that MSI is considered a favorable prognostic factor and a strong negative predictive factor for 5-fluorouracil therapy in stage II colon cancer [6]. However, in cases of metastasis, the impact of MSI on survival is still controversial, with conflicting results reported in studies within this area.

In summary, our investigation has underscored the importance of *KMT2A* mutations and dMMR mutational signatures as independent prognostic markers in mCRC. However, we acknowledge the limitations inherent in our study, particularly regarding the size of our cohort. The possibility of data overinterpretation cannot be overlooked, and we emphasize the necessity of further validation within larger and independent patient cohorts to solidify these preliminary insights.

Supplementary Material

Supplementary figures and tables.

<https://www.jcancer.org/v15p3140s1.pdf>

Acknowledgements

We would like to acknowledge all patients and their families for their contributions to this study.

Funding

This project was supported by the National Natural Science Foundation of China (grant nos. 82273407), the Natural Science Foundation of Jiangsu Province (BK20201495) and Jiangsu Province Capability Improvement Project through Science, Technology and Education (CXZX202204).

Data availability

The datasets utilized or scrutinized during the present investigation are obtainable from the corresponding author upon a reasonable entreaty.

Ethics declaration

Ethics Committee of The First Affiliated Hospital of Nanjing Medical University (2022-SR-294)

Author contributions

HZH, CJS, and WYZ conceived and devised the study. DQL, JKS, WYG and JLY oversaw patient information management and sample collection. HZH, CJS, and FFW conducted data analysis and interpretation. HZH and DQL authored and refined the manuscript. The study was under the supervision of LJZ and XDY. All authors contributed to the article and endorsed the submitted version.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021; 71: 209-49.
2. Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A, et al. Cancer treatment and survivorship statistics, 2022. *CA Cancer J Clin.* 2022; 72: 409-36.
3. Johnston PG. Stage II colorectal cancer: to treat or not to treat. *Oncologist.* 2005; 10: 332-4.
4. Jin Z, Sanhueza CT, Johnson B, Nagorney DM, Larson DW, Mara KC, et al. Outcome of Mismatch Repair-Deficient Metastatic Colorectal Cancer: The Mayo Clinic Experience. *Oncologist.* 2018; 23: 1083-91.
5. Goldstein J, Tran B, Ensor J, Gibbs P, Wong HL, Wong SF, et al. Multicenter retrospective analysis of metastatic colorectal cancer (CRC) with high-level microsatellite instability (MSI-H). *Ann Oncol.* 2014; 25: 1032-8.
6. Sargent DJ, Marsoni S, Monges G, Thibodeau SN, Labianca R, Hamilton SR, et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J Clin Oncol.* 2010; 28: 3219-26.
7. Modest DP, Ricard I, Heinemann V, Hegewisch-Becker S, Schmiegel W, Porschen R, et al. Outcome according to KRAS-, NRAS- and BRAF-mutation as well as KRAS mutation variants: pooled analysis of five randomized trials in metastatic colorectal cancer by the AIO colorectal cancer study group. *Ann Oncol.* 2016; 27: 1746-53.
8. Rashid A, Shen L, Morris JS, Issa JP, Hamilton SR. CpG island methylation in colorectal adenomas. *Am J Pathol.* 2001; 159: 1129-35.
9. Shiovitz S, Bertagnolli MM, Renfro LA, Nam E, Foster NR, Dzieciatkowski S, et al. CpG island methylator phenotype is associated with response to adjuvant irinotecan-based therapy for stage III colon cancer. *Gastroenterology.* 2014; 147: 637-45.

10. Juo YY, Johnston FM, Zhang DY, Juo HH, Wang H, Pappou EP, et al. Prognostic value of CpG island methylator phenotype among colorectal cancer patients: a systematic review and meta-analysis. *Ann Oncol.* 2014; 25: 2314-27.
11. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30: 2114-20.
12. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25: 1754-60.
13. Tarasov A, Vilella AJ, Cuppen E, Nijman JJ, Prins P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics.* 2015; 31: 2032-4.
14. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20: 1297-303.
15. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012; 22: 568-76.
16. Newman AM, Bratman SV, Stehr H, Lee LJ, Liu CL, Diehn M, et al. FACTERA: a practical method for the discovery of genomic rearrangements at breakpoint resolution. *Bioinformatics.* 2014; 30: 3390-3.
17. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol.* 2016; 12: e1004873.
18. Wang S, Li H, Song M, Tao Z, Wu T, He Z, et al. Copy number signature analysis tool and its application in prostate cancer reveals distinct mutational processes and clinical outcomes. *PLoS Genet.* 2021; 17: e1009557.
19. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* 2013; 3: 246-59.
20. Wang F, Cheng H, Zhang X, Shan L, Bai B, Chen K, et al. Comparative genomic signatures in young and old Chinese patients with colorectal cancer. *Cancer Med.* 2021; 10: 4375-86.
21. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell.* 2018; 173: 321-37.e10.
22. Knijnenburg TA, Wang L, Zimmermann MT, Chambwe N, Gao GF, Cherniack AD, et al. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Rep.* 2018; 23: 239-54.e6.
23. Nguyen B, Fong C, Luthra A, Smith SA, DiNatale RG, Nandakumar S, et al. Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. *Cell.* 2022; 185: 563-75.e11.
24. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer.* 2007; 7: 823-33.
25. Liedtke M, Cleary ML. Therapeutic targeting of MLL. *Blood.* 2009; 113: 6061-8.
26. Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* 2013; 502: 333-9.
27. Wan C, Mahara S, Sun C, Doan A, Chua HK, Xu D, et al. Genome-scale CRISPR-Cas9 screen of Wnt/ β -catenin signaling identifies therapeutic targets for colorectal cancer. *Sci Adv.* 2021; 7.
28. Fang Y, Zhang D, Hu T, Zhao H, Zhao X, Lou Z, et al. KMT2A histone methyltransferase contributes to colorectal cancer development by promoting cathepsin Z transcriptional activation. *Cancer Med.* 2019; 8: 3544-52.
29. Liao C, Huang W, Lin M, Li H, Zhang Z, Zhang X, et al. Correlation of KMT2 family mutations with molecular characteristics and prognosis in colorectal cancer. *Int J Biol Markers.* 2022; 37: 149-57.
30. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet.* 2003; 72: 1117-30.
31. Chen S, Iversen ES, Friebe T, Finkelstein D, Weber BL, Eisen A, et al. Characterization of BRCA1 and BRCA2 mutations in a large United States sample. *J Clin Oncol.* 2006; 24: 863-71.
32. Easton DF, Ford D, Bishop DT. Breast and ovarian cancer incidence in BRCA1-mutation carriers. *Breast Cancer Linkage Consortium. Am J Hum Genet.* 1995; 56: 265-71.
33. Cancer risks in BRCA2 mutation carriers. *J Natl Cancer Inst.* 1999; 91: 1310-6.
34. Thompson D, Easton DF. Cancer Incidence in BRCA1 mutation carriers. *J Natl Cancer Inst.* 2002; 94: 1358-65.
35. Oh M, McBride A, Yun S, Bhattacharjee S, Slack M, Martin JR, et al. BRCA1 and BRCA2 Gene Mutations and Colorectal Cancer Risk: Systematic Review and Meta-analysis. *J Natl Cancer Inst.* 2018; 110: 1178-89.
36. Lin Z, Liu J, Peng L, Zhang D, Jin M, Wang J, et al. Complete pathological response following neoadjuvant FOLFOX chemotherapy in BRCA2-mutant locally advanced rectal cancer: a case report. *BMC Cancer.* 2018; 18: 1253.
37. Soyano AE, Baldeo C, Kasi PM. BRCA Mutation and Its Association With Colorectal Cancer. *Clin Colorectal Cancer.* 2018; 17: e647-e50.
38. King MC. "The race" to clone BRCA1. *Science.* 2014; 343: 1462-5.
39. Moynahan ME, Jasin M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol.* 2010; 11: 196-207.
40. Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature.* 2005; 434: 917-21.
41. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature.* 2005; 434: 913-7.
42. Koh G, Degasperi A, Zou X, Momen S, Nik-Zainal S. Mutational signatures: emerging concepts, caveats and clinical applications. *Nat Rev Cancer.* 2021; 21: 619-37.
43. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature.* 2013; 500: 415-21.
44. Kim J, Mouw KW, Polak P, Braunstein LZ, Kamburov A, Kwiatkowski DJ, et al. Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. *Nat Genet.* 2016; 48: 600-6.
45. Kunkel TA, Erie DA. Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annu Rev Genet.* 2015; 49: 291-313.
46. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; 144: 646-74.
47. Macintyre G, Goranova TE, De Silva D, Ennis D, Piskorz AM, Eldridge M, et al. Copy number signatures and mutational processes in ovarian carcinoma. *Nat Genet.* 2018; 50: 1262-70.
48. Grilley M, Holmes J, Yashar B, Modrich P. Mechanisms of DNA-mismatch correction. *Mutat Res.* 1990; 236: 253-67.
49. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012; 487: 330-7.
50. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology.* 2010; 138: 2044-58.