Am J Cancer Res 2021;11(12):6119-6132
www.ajcr.us /ISSN:2156-6976/ajcr0137189

Original Article

Automated assessment of DNA ploidy, chromatin organization, and stroma fraction to predict prognosis and adjuvant therapy response in patients with stage II colorectal carcinoma

Zhixun Zhao[1*], Xiaochen Zhang[2*], Zhongwu Li[3*], Yibo Gao[4*], Xu Guan[1], Zheng Jiang[1], Zheng Liu[1], Ming Yang[1], Haipeng Chen[1], Xiaolong Ma[1], Runkun Yang[5], Zhao Lu[1], Hengchang Liu[1], Lujing Yang[3], Aiwen Wu[6], Shuangmei Zou[7], Xishan Wang[1,2]

[1]Department of Colorectal Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [2]Department of Gastroenterology and Hepatology, Institute of Clinical Medicine, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan; [3]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Department of Pathology, Peking University Cancer Hospital & Institute, Beijing, China; [4]Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [5]Department of Colorectal Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, China; [6]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Department of Colorectal Surgery, Peking University Cancer Hospital & Institute, Beijing, China; [7]Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. [*]Equal contributors.

Received July 9, 2021; Accepted October 3, 2021; Epub December 15, 2021; Published December 30, 2021

Abstract: DNA ploidy, tumor stroma, and chromatin organization have important implications in tumorigenesis and patient outcome. Automated image cytometry tools were developed to quantitatively measure DNA ploidy (P), stroma fraction (S), and chromatin organization or Nucleotyping (N). This study aimed to discover their clinical value in different stages of colorectal cancer (CRC) in a Chinese patient population. A total of 496 CRC patients of stages I, II, and LMCRC (liver metastatic CRC) were enrolled in this study. Stage II CRC patients with diploidy, low-stroma, or chromatin homogenous status predicted significantly higher 5-year OS and DFS. We constructed a PSN-panel enabled the stage II patients to be further stratified into low-, middle-, high-risk groups, the 5-year OS (89.5% *vs* 67.9% *vs* 60.9%, P<0.001) and DFS (86.0% *vs* 62.3% *vs* 53.6%, P<0.001) were stratified significantly. In addition, when combined the PSN-panel with T stage or MSS status in stage II patients, the PSN-low risk patients showed significant longer 5-year OS and DFS than the PSN-high risk patients in T3 (OS: 86.3% *vs* 65.3%, P=0.015; DFS: 83.5 *vs* 59.8%, P=0.013) or MSS (OS: 86.4% *vs* 63.9%, P=0.005; DFS: 85.5 *vs* 57.8%, P=0.003) patients. Finally, in the group of stage II patients with at least one high-risk factor (non-diploidy, high-stroma, chromatin heterogenous), patients who received adjuvant therapy showed significantly longer OS (72.1% *vs* 48.3%, P=0.007) and DFS (64.5% *vs* 43.9%, P=0.015) than those who did not receive adjuvant therapy. In contrast, P, S, N couldn't predict the prognosis of stage I and LMCRC patients. Overall, our data demonstrate that the PSN panel is an accurate prognostic tool that can guide treatment decisions for Chinese stage II CRC patients.

Keywords: Colorectal cancer, ploidy, stroma-tumor fraction, Nucleotyping, prognosis, adjuvant therapy

Introduction

Colorectal cancer (CRC) is one of the most common malignant gastrointestinal tumors with high incidence and mortality rates. In China, it is estimated that there are over 376,000 new CRC cases and 191,000 attributable deaths every year [1]. Pathological features are important factors to guide clinical decisions, but patients with identical pathological staging

and receiving similar treatment may experience different clinical outcomes [2]. Especially in stage II CRC patients, the postoperative treatment decision is based on the Microsatellite Instability (MSI) status and clinical high-risk factors. However, in a previous study, no significant benefits from adjuvant chemotherapy were observed with regard to RFS or DSS in 1286 high-risk stage II CRC patients [3]. Therefore, it would be critical to developing complementary biomarkers to more accurately stratify patients who may benefit from certain clinical treatment schemes and improve patient outcomes.

The DNA-ploidy status of tumor tissue is a useful prognostic indicator of disease progression risk in many tumors [4]. Danielsen *et al*. developed an automatic tool, which could automatically group the nucleus into different galleries, and analyze the DNA ploidy status of epithelial cells based on more than 1000 cell nucleus. The prognostic value of ploidy status detected by this method was demonstrated to be significantly stronger than that associated with MSI status in CRC stage II patients [5].

Another biomarker is the stroma-tumor fraction that measures the proportion of surrounding stroma relative to tumor tissue based on HE stained tissue sections. A series of older studies have identified that a high stroma ratio is associated with a high risk of recurrent disease and poorer survival in colorectal cancers [6, 7], especially in CRC stage II and III [8]. The traditional tumor-interstitial ratio analysis is performed by a pathologist under a microscope and the results are relatively rough. Later on, David *et al*. developed an automated analysis tool to quantify the stroma ratio [5]. The combination of automated ploidy and stroma analyses proved to be useful in providing an accurate prognosis in stage II CRC patients [5].

Meanwhile, the high mutation frequency in tumor cell DNA [9] and epigenetic modifications that occur during tumor development [10] would likely be reflected in the disruption of chromatin organization. Based on this theory, Kleppe and colleagues [11] developed an automated machine learning algorithm to determine chromatin organization based on tumor cell nuclei image texture analysis, named Nucleotyping. The Nucleotyping method was demonstrated to be a powerful pan-cancer prognostic tool and in direct comparison to MSI status in CRC patients, was shown to more accurately predict patient survival [11].

The biomarkers of DNA ploidy, stroma-tumor ratio, and Nucleotyping have been clinically validated to predict patient survival in stage II colorectal cancer within a European patient population [5, 11], but the DNA ploidy, stroma could not predict the prognosis of Stage III CRC patients (HR=1.43, P=0.14) [5]. Recently, a Chinese study proved that these three markers could predict the prognosis of patients with high-risk pathological stage II colon cancer [12]. However, the role of these three markers in stage I and IV CRC patients is still unclear.

In this study, we sought to validate the predictive value of DNA-ploidy status (P), the stroma-tumor fraction (S), and Nucleotyping (N) in CRC patients of stage I, II, and IV from a major Chinese hospital. Firstly, we will verify the prognostic value of these three biomarkers in stage II colorectal patients and the relationship between them and the efficacy of postoperative adjuvant therapy. On the other hand, the 5-year survival rate for stage I CRC patients can reach as high as 91% [13], there were still some patients with poor prognosis. We tried to explore whether P, S, and N can help screen stage I patients with poor prognosis. Lastly, the 5-year survival rate of stage IV colorectal cancer is only 12%, and the incidence of liver metastasis is high [13]. This research analyzes the status of P, S, and N in CRC patients with liver metastasis, providing a basis for exploring biomarkers related to liver metastasis.

Material and methods

Patients

The specimens in this study were collected from the patients who underwent the surgical resection between 2006 and 2012 at Cancer Institute & Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College. Patient diagnoses were confirmed according to the 7th edition AJCC TNM Classification criteria. All cases had complete clinical information including age, gender, tumor location, gross pathological type, histological grading, TNM classification, adjuvant therapy, regular follow-up information until December 31st, 2017. This study was approved by the Clinical

Research Ethics Committee of Cancer Institute & Hospital, Chinese Academy of Medical Sciences. The inclusion criteria of the BJCH cohort were shown in the Supplementary Information.

There were 143 patients in stage I and 179 patients in stage II, and these cases did not receive systemic neoadjuvant chemotherapy before the surgery. In addition, the primary tumor specimens from 174 stage IV CRC patients with liver metastatic (LMCRC) were used to establish another LMCRC group, out of which 27 (15.52%) patients received neoadjuvant therapy before surgery.

Tumor tissue sampling

The pathologist selected one tumor block deemed representative from each patient and a 5-μm section stained with H&E was used to define the tumor region and analyze the stroma fraction. For Ploidy and Nucleotyping analysis, one or two 50-μm sections containing more than 90% representative tumor tissue marked by a pathologist were cut from paraffin-embedded blocks. Nuclear monolayers were prepared from 50-μm sections, and stained by the Feulgen method as previously described [14].

Measurement of DNA ploidy

Tumor DNA ploidy analysis by image cytometry was performed with DNA Ploidy Working Station (PWS, Room4, Kent, UK) according to the previously reported technique [15]. Briefly, images of Feulgen-stained nuclei were captured by a high-resolution digital scanner (Aperio AT2, Leica, Germany), and the nuclear images were automatically grouped into different galleries for tumor nuclei, reference nuclei, and discarded nuclei by the PWS classifier. DNA ploidy histograms were created from the integrated optical density (IOD) of the tumor nucleus and reference nucleus. The reference nucleus was used as an internal diploid control, and DNA ploidy of tumor nucleus was classified into three groups: diploidy, aneuploidy and tetraploidy, according to the previous report [16].

Nucleotyping analysis

The Nucleotyping tool was developed based on the machine learning algorithm, and it could automatically assess the chromatin organization [11]. Firstly, the image of the tumor nucleus was imported into the PWS Classifier (Room 4, Kent, UK), and then the nucleus was grouped into 11 groups according to the areas of the nucleus. Then the chromatin organization was quantified by computing the entropy of pixel grey levels in a subregion of a nucleus. The frequency in which each pair of entropy and center grey levels occur throughout a nucleus was stored in a two-way table, known as the grey level entropy matrix (GLEM). GLEMs stratified on the nuclear area and subregion size was concatenated to form a four-dimensional expansion of the GLEM called GLEM4D. In a previous study [11], an adaptive machine-learning algorithm was applied to quantify the association between each element of the GLEM4D and the outcome of the patient. In the current study, these pre-trained weights were directly applied to predict the outcome of a patient based on the GLEM4D representation of its tumor. The result is a continuous value termed the chromatin value, which describes the overall amount of chromatin disorder in a given patient sample. As determined in the previous study [11], the chromatin values less than 0.044 were labeled as heterogeneous, while those with chromatin values higher than or equal to 0.044 were labeled homogeneous.

Stroma-tumor fraction

The stroma-tumor fraction was automatically calculated by the software of Stroma Analyzer (Room 4, Kent, UK) [5]. The H&E histological images were scanned with the 40× lens selection on an Aperio AT2 digital slide scanner (Leica, Germany). The tumor regions in the H&E images were annotated by a pathologist. The tumor area and the stroma area were automatically identified by the software Stroma Analyzer. Then the stroma fraction was calculated according to the ratio of stroma area to the area of the annotated regions. Tumors with stroma fraction less or equal to 50% were labeled low stroma, while those with stroma fraction greater than 50% were labeled high stroma [5, 8].

Statistical analysis

The endpoints were overall survival (OS) for stage I, OS, and disease-free survival (DFS) for stage II, and progression-free survival (PFS) for LMCRC. OS is defined as the days from the date

Table 1. Demographic and clinical characteristic of patients in the CAMCH cohort

Variables	Stage I N (%)	Stage II N (%)	Stage IV N (%)
Age, years			
Mean ± SD	60.49±9.99	58.54±11.05	57.48±10.84
Range	23-78	30-77	21-81
Gender			
Male	77 (53.8)	109 (60.9)	107 (61.5)
Female	66 (46.2)	70 (39.1)	67 (38.5)
Tumor site			
Colon	38 (26.6)	91 (50.8)	97 (55.7)
Rectum	105 (73.3)	88 (49.2)	77 (44.3)
Histological grade			
High	38 (26.6)	21 (11.7)	3 (1.7)
Middle	100 (69.9)	149 (83.2)	25 (14.4)
Low	5 (3.5)	9 (5.1)	146 (83.9)
Lymph nodes sampling			
≥12	99 (69.2)	179 (100)	174 (100)
<12	44 (30.8)		
pT stage			
pT1	143 (100)		
pT2			2 (1.1)
pT3		156 (87.2)	100 (57.5)
pT4		23 (12.8)	72 (41.4)
Adjuvant therapy			
YES	-	112 (62.6)	174
No	-	67 (37.4)	0
Mismatch repair status			
MSI-H	3 (2.1)	16 (8.9)	17 (9.8)
MSS/MSS-L	140 (97.9)	163 (91.1)	157 (90.2)
Ploidy			
Diploid	39 (27.3)	73 (40.8)	45 (25.9)
Non-diploid	104 (72.7)	106 (59.2)	129 (74.1)
Stroma			
Low-stroma fraction	140 (97.9)	136 (76.0)	107 (61.5)
High-stroma fraction	3 (2.1)	43 (24.0)	67 (38.5)
Nucleotyping			
Chromatin heterogeneous	38 (26.6)	54 (30.2)	47 (27)
Chromatin homogeneous	105 (73.4)	125 (69.8)	127 (73)
Total	143	179	174

Note: Stage IV patients only included the liver metastasis Colorectal cancer patients.

of the first surgery to the date of death for any reason or the date of the last follow-up. DFS is defined as the days from the date of the first operation to the date of death for any cause or first local recurrence or metastasis.

Correlation between ploidy and Nucleotyping was analyzed using a cross-tab chi-square test. Kaplan-Meier survival curves were plotted with log-rank tests to compare the OS and DFS. Univariate and multivariate analysis was undertaken using Cox proportional hazards regression model. The clinically relevant variables of T stage, MSI-status, and adjuvant therapy were included in the multivariate analysis. The correction for multiple comparisons was performed and $P<0.05$ was considered significant for each multivariable model. The calculations were performed with R version 3.6.1. A value of $P<0.05$ was considered significant.

Results

Demographic and clinical characteristic

The patient clinical characteristics are summarized in **Table 1**. The median age at the time of surgery was 58 years (range 21-82 years), a male preponderance (59.1% vs 40.9%), a higher proportion of rectal cancer patients (54.4% vs 45.6% colon cancer), a majority of patients (91.1%) had ≥12 lymph nodes samplings. Stage I patients were all T1 (100%), most stage II patients were T3 (87.2% vs 12.8% T4), only 2 patients were T2 in stage IV (1.1% vs 57.5% T3 and 41.4% T4). None of the stage I patients received adjuvant therapy, 62.6% Stage II, and all of the stage IV patients received adjuvant therapy.

The fraction of MSI-H patients in stage I, II, and LMCRC stages were 2.1%, 8.9%, and 9.8%,

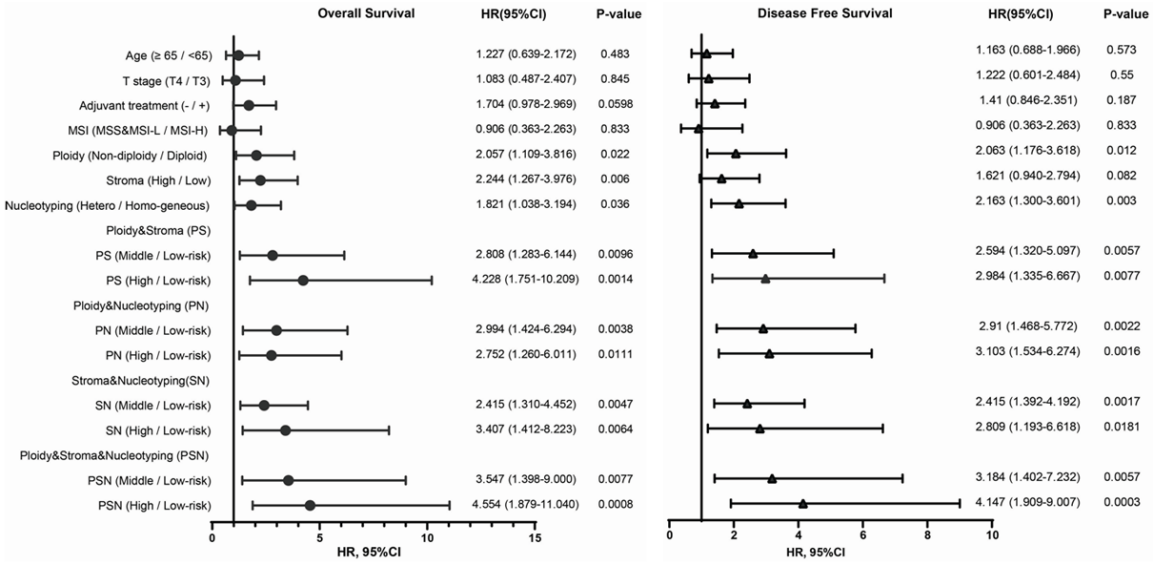


Figure 1. Forest plot of potential prognostic factors for stage II CRC patients, including tumor Ploidy, Stroma and Nucleotyping status.

respectively. The highest proportion of diploid patients was observed in Stage II patients (40.8% vs 27.3% in stage I and 25.9% in stage IV). The proportion of patients with a high-stroma fraction increased with the higher staging groups (2.1% in Stage I vs 24% in Stage II vs 38.5% in Stage IV). The proportion of chromatin heterogeneity, as measured by Nucletyping, was similar in the three stages.

Prognostic significance of ploidy, stroma-tumor fraction, and Nucleotyping for stage II CRC patients

Firstly, we verified the prognostic value of Ploidy, Stroma, and Nucleotyping for OS and DFS in stage II CRC (n=179). The median follow-up was over 59 months, during which 50 patients died, and 60 patients relapsed. Univariate analysis of ploidy (*P*=0.022; HR= 2.057, [95% CI: 1.109-3.816]), stroma-tumor fraction (*P*=0.006; HR=2.244, [95% CI: 1.267-3.976]), and Nucleotyping (*P*=0.036; HR= 1.821, [95% CI: 1.038-3.194]) showed that all three are significant prognostic markers for OS in stage II patients (**Figure 1** and Table S1). The 5y OS of patients with non-diploid, high-stroma-fraction, and chromatin heterogenous was 14.8%, 21.4%, and 15.7% lower than the diploidy, low-stroma-fraction, and chromatin homogenous patients, respectively (Figure S1 and Table S1).

When applying univariate analysis for DFS, the 5y DFS of patients with non-diploid, high-stroma-fraction, and chromatin heterogenous was 17.3%, 14.1%, and 23.6% lower. Patients with either non-diploid tumors (*P*=0.012; HR=2.063, [95% CI: 1.176-3.618]) or chromatin heterogeneous status (*P*=0.003; HR=2.163, [95% CI: 1.300-3.601]) was associated with higher mortality risk. However, stroma-tumor rate did not have a significant predictive value in DFS (High stroma vs Low stroma: *P*=0.082; HR=1.621, [95% CI: 0.940-2.794]). The results were shown in **Figure 1** and Table S1.

We then performed the multivariate analysis with Cox regression. Besides P, S, N, our analysis included the clinical factors of T stage, MSI-status, and adjuvant therapy as these were key clinical factors for stage II patients. Ploidy and Nucleotyping were confirmed as independent prognostic factors for both OS and DFS (**Table 2**).

The prognostic value of the combination of ploidy, stroma-tumor fraction, and Nucleotyping in stage II CRC patients

A previous study demonstrated that combining ploidy-stroma marker improved prognostic accuracy among stage II tumors [5]. Hence, we tested the prognostic accuracy of different combinations of ploidy, stroma-tumor fraction, and Nucleotyping in our stage II patient cohort

Table 2. Multivariate analysis of Ploidy, Stroma and Nucleotyping as standalone or combined factors as predictors of OS and DFS in Stage II CRC patients

Independent variables	OS		DFS	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Ploidy	2.099 (1.109-3.973)	0.023	2.188 (1.222-3.918)	0.008
Nucleotyping	1.922 (1.089-3.394)	0.024	2.273 (1.358-3.804)	0.002
Stroma	2.33 (1.313-4.137)	0.004	1.668 (0.966-2.88)	0.067
PN	1.618 (1.137-2.302)	0.007	1.738 (1.26-2.396)	0.001
PS	1.999 (1.337-2.99)	0.001	1.747 (1.213-2.515)	0.003
PNS	2.371 (1.467-3.832)	<0.001	2.195 (1.418-3.397)	<0.001

All variables were separately adjusted with T stage, MSI status and adjuvant treatment; A two-sided *P*-value of less than 0.05 was considered statistically significant.

(**Figure 1** and Table S1). The high stroma ratio, non-diploidy, and chromatin heterogeneous were the high-risk factors. In the combined PS-panel the patients were divided into the low-risk, middle-risk, and high-risk groups, which contained zero, one/two, and three high-risk factors, respectively. Patients classified as PS-low-risk group had a better OS, compared to the PS-middle-risk group (HR=2.808, [95% CI: 1.283-6.144]) and PS-high-risk group (HR= 4.228, [95% CI: 1.751-10.209]).

The Ploidy and Nucleotyping biomarkers indicate the changes in DNA content or organization, respectively. So, the correlations between them were analyzed. In the 157 patients with diploid tissue, 150 patients were with chromosome homogeneous tumors. And in the 139 patients with chromatin heterogeneous tissue, 132 patients were non-diploidy (Table S2). The combination of Ploidy and Nucleotyping may improve the value in prognostic stratification. The patients were divided into three groups in the PN panel as the PS panel. The PN-low risk group showed a better OS compared to the PN-middle-risk group (HR=2.994, [95% CI: 1.424-6.294]) and PS-high-risk group (HR= 2.752, [95% CI: 1.260-6.011]). At the same time, the above panels also showed good prognostic values when using DFS as the endpoint (Table S1 and Figure S2). For the PS-panel showed a good group trend and the association of Ploidy and Nucleotyping, the Nucleotyping and Stroma were also used to construct a model. The SN-panel could stratify the OS and DFS of the patients (Table S1 and Figure S2).

When the three biomarkers were combined into a single panel (PSN), the prognostic value became even stronger than the combination of the two factors panel. Based on the PSN panel, patients were segregated into the PSN-low-risk, PSN-middle-risk, and PSN-high-risk groups, which contained zero, one/two, and three high-risk factors, respectively. Importantly, the PSN-middle-risk and PSN-high-risk groups both showed additive mortality risk for both OS (PSN-middle-risk: HR=3.547, [95% CI: 1.398-9.000]; PSN-high-risk group: HR=4.554, [95% CI: 1.879-11.040]) and DFS (PSN-middle-risk: HR=3.184, [95% CI: 1.402-7.232]; PSN-high-risk group: HR=4.147, [95% CI: 1.909-9.007]) when compared to PSN-low-risk patients (Table S1 and **Figure 2**). In the multivariable cox model, the PSN was the dominant contributory factor on DFS, patients in the PSN-high-risk group may have the worse OS and DFS compared with the PSN-low-risk group (OS: HR= 2.371, [95% CI: 1.467-3.832], *P*<0.001; DFS: HR=2.195, [95% CI: 1.418-3.397], *P*<0.001) (**Table 2**).

To further validate the prognostic significance of this PSN panel, we carried out the survival analysis on patient samples collected from BJCH-cohort (Table S3). As shown in Figure S3, the PSN panel could separate PSN-high-risk patients from the other low- and middle-risk patients (PSN-middle-risk: HR=1.099, [95% CI: 0.386-3.135]; PSN-high-risk group: HR=3.551, [95% CI: 1.411-8.938]) when using DFS as an endpoint.

The PSN could stratify the prognosis of stage II patients with T3 or MSS/MSI-L status

T staging is an important guiding factor for postoperative adjuvant treatment of patients with stage II colorectal cancer. According to NCCN guidelines [17], stage II CRC patients

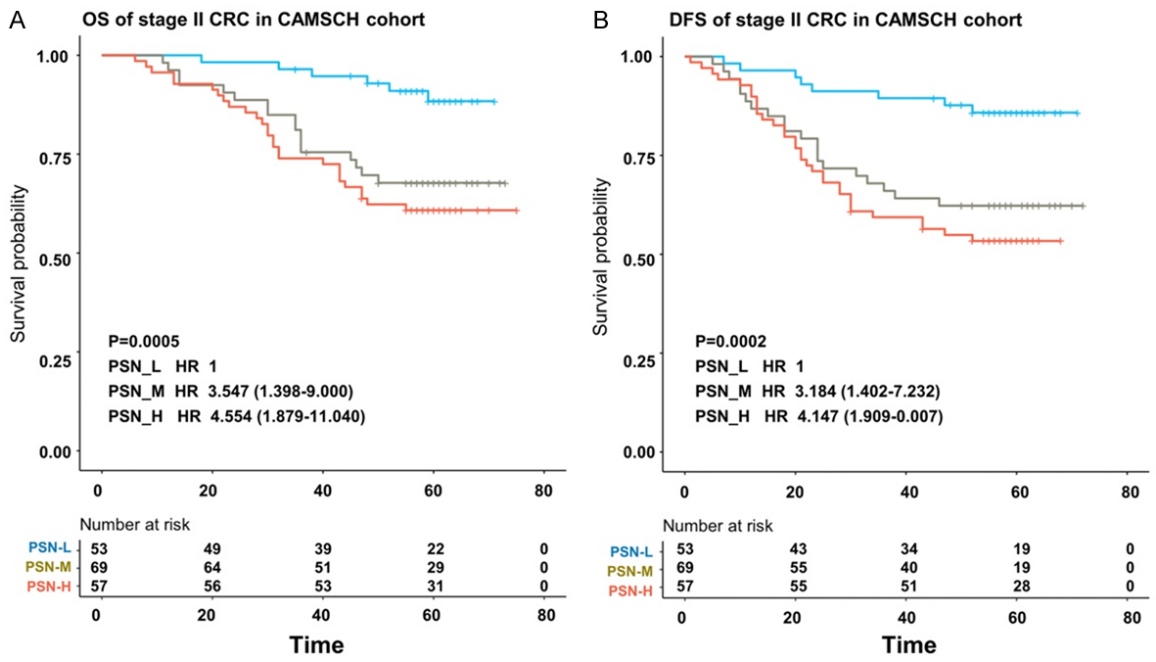


Figure 2. The prognostic significance of PNS panel in CAMSCH cohort. Kaplan-Meier plots illustrating (A) overall survival (OS) and (B) disease free survival (DFS) of CAMSCH cohort stage II CRC patient with tumors classified according to the Ploidy-Nucleotyping-Stroma (PSN) panel.

with T4 tumors are classified into the high-risk group that should receive adjuvant chemotherapy. However, in our patient cohort, the 5-year OS and 5-year DFS between the T3 and T4 patients showed no statistically significant difference (Table S1). As the PSN could stratify the prognosis of stage II CRC patients, we investigated whether the combination of PSN and T stage could further stratify the patients' risk. The results showed that the T3 patients with PSN high-risk factors, the 5-year OS (65.3%) and DFS (59.8%) of patients are close to those of T4 patients (OS, 69.6%; DFS 60.6%). The 5y OS (86.3%) and DFS (83.5%) of low-risk T3-PSN patients were significantly higher than those of T4 patients (**Figure 3A**, **3B** and Table S4).

The NCCN guidelines also indicate that patients with MSI-H have a better prognosis compared to MSS/MSI-L patients [17]. This study combined microsatellite instable status and PSN factors to analyze the OS and DFS of patients. The results found that the PSN factors could stratify the prognosis of MSS/MSI-L patients. The MSS/MSI-L combining with PSN-low-risk factor patients, 5-year OS (86.4%) and 5-year DFS (85.5%) were even better than MSI-H patients (OS, 81.3%; DFS 68.8%), while MSS/MSI-L & PSN high-risk patients had the worst OS (63.9%) and DFS (57.8%) (**Figure 3C**, **3D** and Table S4). Overall, these results demonstrate that the PSN panel may have a superior prognostic value compared to either T staging or MSI status.

The PSN-panel could predict the efficacy of postoperative adjuvant therapy in stage II CRC patients

In stage II CRC patients, 67 did not receive adjuvant therapy and 111 received postoperative adjuvant therapy. Adjuvant therapy did not improve the patient's 5 years OS (HR=1.702, 95% CI: 0.977-2.966, P=0.06) or 5 years DFS (HR=1.407, 95% CI: 0.844-2.345, P=0.190) (**Table 3** and Figure S4).

We analyzed whether the three markers of P, S, and N could predict the efficacy of adjuvant therapy. As shown in **Table 3**, patients with nondiploid (Figure S5) or chromatin heterogeneity (Figure S6) who received adjuvant therapy had higher 5y OS and DFS than patients who did not receive adjuvant therapy. The 5y OS improved 21.4% (P=0.022) in nondiploid patients and 24.9% (P=0.05) in chromatin heterogeneity patients.

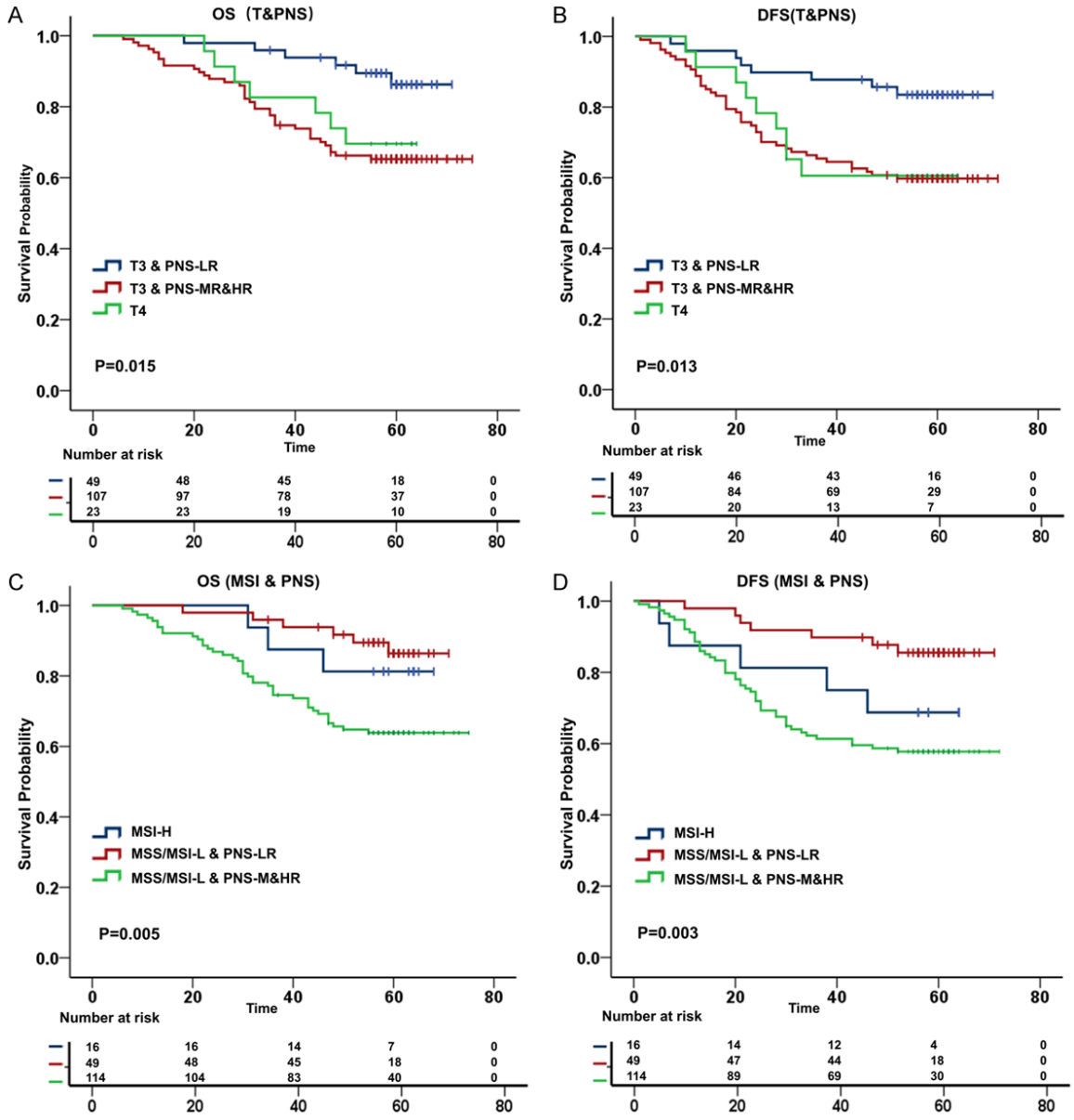


Figure 3. Kaplan-Meier plots illustrating the overall survival (OS) and disease free survival (DFS) of stage II CRC patients when classified by the combination of Ploidy-Nucleotyping-Stroma (PNS) panel status with T stage (A, B) or microsatellite instability status (C, D).

We further analyzed the effect of adjuvant therapy on OS and DFS in patients when stratified by the combined PSN panel. For patients with low risk factors (PSN-L), adjuvant therapy did not improve OS (HR=0.669, [95% CI: 0.123-3.657], P=0.643) or DFS (HR=0.435, [95% CI: 0.088-2.153], P=0.307), and may even be detrimental for these patients (**Figure 4A**, **4C**). However, for patients with at least one high-risk factor (PSN-M & PSN-H), adjuvant therapy significantly improved 5y OS (23.8%) and 5y DFS (20.6%), while the patients who did not receive adjuvant therapy would have worse OS (HR=2.252, [95% CI: 1.245-4.071], P=0.007) and DFS (HR=1.973, [95% CI: 1.142-3.407], P=0.015).

Prognostic significance of ploidy, stroma-tumor fraction, and Nucleotyping for stage I CRC patients

Next, we analyzed the prognostic value of Ploidy, Stroma, and Nucleotyping in our stage I patient cohort (n=143). We found that all three

Table 3. Kaplan-Meier analysis on overall survival and disease-free survival of stage II CRC patients who recieved (+) or did not receive (-) postoperative adjuvant therapy, classified by T stage, MSI, ploidy, stroma and Nucleotyping status

Variables	Total (-/+)	OS			DFS		
		Events (-/+)	5y OS% (-/+)	*P*-value	Events (-/+)	5y DFS% (-/+)	*P*-value
Adjuvant therapy	67/112	24/26	62.9/76.6	0.06	26/34	61/69.6	0.19
T							
T3	55/101	21/22	59.9/78.1	0.023	21/30	61.7/70.2	0.24
T4	12/11	3/4	75.0/63.6	0.675	5/4	58.3/63.6	0.607
MSI							
MSI-H	8/8	3/0	62.5/	0.063	2/3	75.0/62.5	0.513
MSS	59/104	21/26	62.9/74.8	0.121	24/31	59.1/70.1	0.098
Ploidy (P)							
Diploid	29/44	6/8	76.6/81.4	0.813	6/11	79/74.9	0.634
Non-diploid	38/68	18/18	52.1/73.5	0.022	20/23	47.1/66.1	0.031
Nucleotyping (N)							
Homogenous	49/76	14/15	69.9/80.1	0.268	15/18	69.3/76.3	0.409
Heterogeneous	18/36	10/11	44.4/69.3	0.05	11/16	38.1/55.3	0.098
Stroma (S)							
Low-stroma	53/83	16/15	68.4/81.8	0.092	18/23	66/72.2	0.356
High stroma	14/29	8/11	42.9/62.1	0.234	8/11	40.8/62.1	0.241
PS							
PS-L	26/34	4/4	82/87.9	0.71	4/7	84.4/79.3	0.604
PS-M&H	41/78	20/22	50.7/71.7	0.016	22/27	46/65.3	0.024
PN							
PN-L	27/41	4/6	82.3/85	1	4/8	84.9/80.4	0.594
PN-M&H	40/71	20/20	49.4/71.8	0.015	22/26	44.7/63.3	0.028
PSN							
PSN-L	24/33	2/4	88.8/87.6	0.643	2/6	91.5/81.7	0.307
PSN-M&H	43/79	22/22	48.3/72.1	0.007	24/28	43.9/64.5	0.015

variables had no impact on five years of DFS or OS (Figure S7). As expected in this group, the overall prognosis is better. The median follow-up was 58.5-months, and only ten patients died.

Prognostic significance of ploidy, stroma-tumor fraction, and Nucleotyping for stage IV liver metastatic CRC patients

In the liver metastasis CRC (LMCRC) group (n=174), 123 patients died, and 149 patients relapsed, within a median 34-month follow-up period. Again, Again, not OS or DFS showed significant differences between LMCRC patients classified by ploidy, stroma-tumor fraction, and Nucleotyping characteristics (Figure S1). MSI-H is an important biomarker in stage IV patients and is used to help select the patients who may benefit from PD-1/PD-L1 immunotherapy. Excluding the 17 MSI-H patients in the stage IV patients, the P, S, N factors were used to stratify the prognosis of the MSS/MSI-L patients. Interestingly, it was found that the non-diploidy patients, had worse OS (HR=1.604, [95% CI: 1.030-2.499], P=0.036) and DFS (HR=1.535, [95% CI: 1.022-2.306] in MSS/MSI-L patients, P=0.039) compared with diploidy patients (Figure S8).

Discussions

In the present study, we evaluated the characteristics and prognostic values of three pathological biomarkers in different stages of CRC tumor development. The DNA ploidy, chromatin organization, and stroma-ratio have been reported to have independent prognostic values in many epithelial cancers, including CRC. Aneuploidy in cancer represent an abnormal

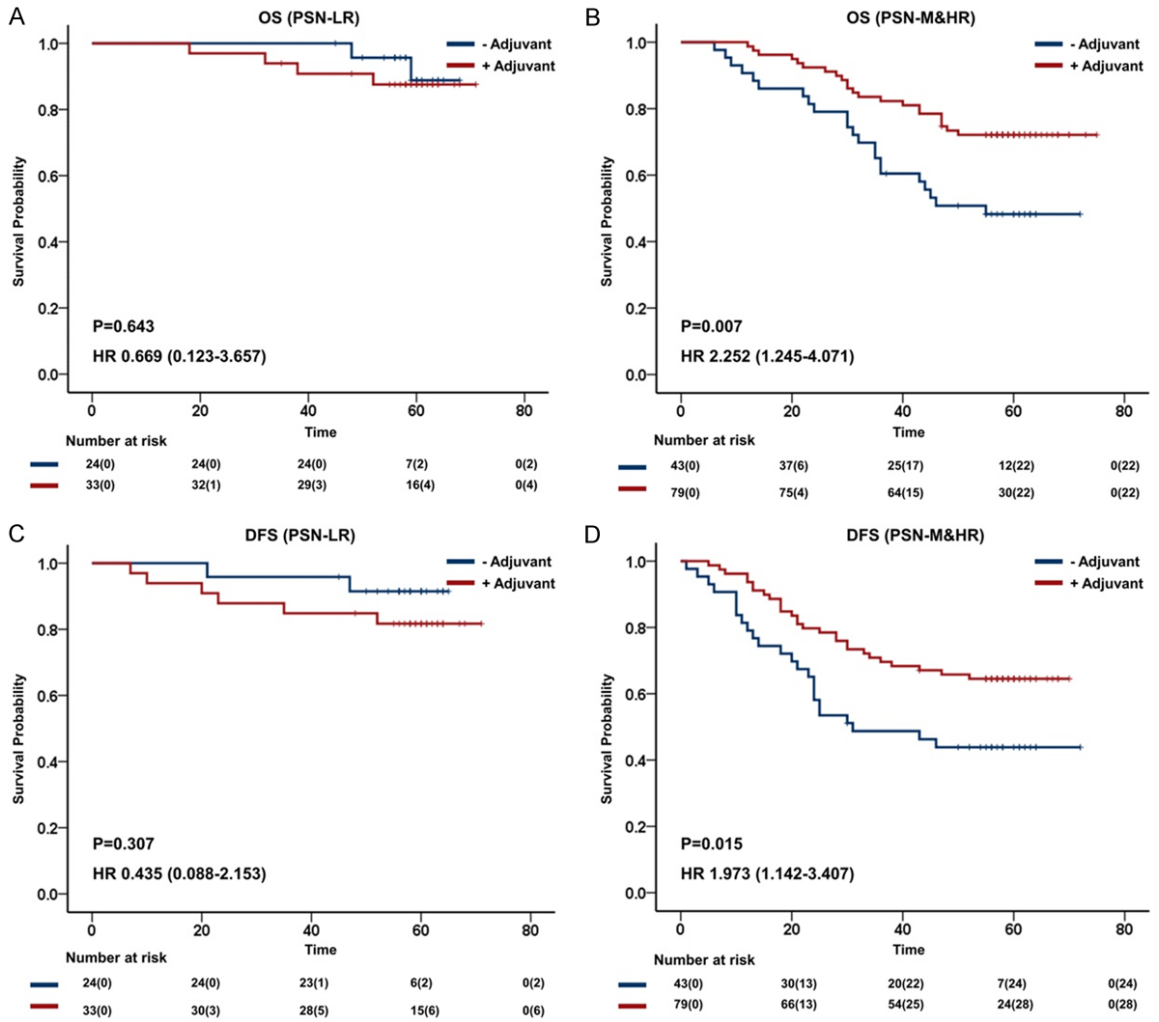


Figure 4. Kaplan-Meier plots illustrating overall survival (OS) and disease-free survival (DFS) in stage II CRC patients who received (+) or did not receive (-) adjuvant therapy, based on Ploidy-Nucletyping-Stroma (PNS) status. (A) OS and (C) DFS in PNS-low-risk patients. (B) OS and (D) DFS in PNS-middle/high-risk patients.

state of cellular DNA content and can be detected by DNA cytometry [15, 16]. Previous studies have established that abnormalities of cellular DNA content were associated with tumorigenesis [18-20]. Vermeulen *et al*. suggested that stroma supplied the tumor with growth factors, cytokines, and metabolites, and stimulated blood vessel formation, which could cause tumorigenesis and induction of epithelial-mesenchymal transition (EMT) [21]. Thus, high stroma-tumor content likely represents the metastatic phenotype of cancer cells. The DNA ploidy/chromatin organization and stroma-tumor ratio were measured by automated digital image analysis of nuclear monolayer [5, 11] and H&E [8] slides respectively. These tools could rapidly and quantitatively detect the changes in DNA level and tissue level. Most previous studies focused on the prognostic significance of these factors in CRC stage II patients [5, 12], the ploidy and stroma were not suitable for prognosis analysis in Stage III CRC [5]. Thus, our study aimed to determine if these biomarkers were broadly applicable to stage I and liver metastatic CRC, and to a Chinese patient population.

We evaluated the three biomarkers as standalone prognostic parameters of stage II CRC patients and found that our results are consistent with previous studies [5, 11, 12]. In contrast, we found no association between 5-years OS and any of the three indexes in stage I patients. Similarly, in LMCRC patients, we

found that there was no association between 5-years PFS and any of the three biomarkers if no prior patient classification was carried out. However, a prior study pointed out that Nucleotyping could stratify the prognosis in stage II patients after MSS/MSI-L or MSI-H classification [11]. In current clinical practice, patients with advanced solid tumors who are MSI-H could accept immunotherapy. However, only 5% of mCRC patients have MSI-H and have a poor prognosis [22]. Considering the importance of MSI status on patient treatment decisions, we wanted to determine whether DNA ploidy and Nucleotyping could further differentiate advanced CRC patients who were MSS or MSI-L and assist in the treatment decision for these patients. Specifically, we demonstrated that ploidy classification could predict the prognosis of LMCRC patients after MSS/MSI-L classification (Figure S2). A recent study has shown that melanoma patients with diploidy had longer survival after immunotherapy [23]. And we found that ploidy status can stratify the survival of MSS patients in this study. Ploidy may be a candidate marker to select MSS patients who could benefit from immunotherapy.

While both ploidy and Nucleotyping reflect changes in DNA, the ploidy indicates the DNA content and Nucleotyping indicates the changes in chromatin organization. Our data showed that chromatin heterogeneous patients are more likely to have non-diploid phenotypes, and diploid patients are also more likely to be chromatin homogenous, suggesting that both biomarkers correlated on a cellular level. According to this result, we attempted the combination of these two markers. The PN model could better stratify the OS and DFS, and the prognosis of patients in the high-risk group was worst.

In a previous study, the combination of ploidy and stroma could predict the prognosis in stage II CRC patients [5]. As recently reported that the PS panel and the Nucleotyping could predict the prognosis of stage II colon cancer patients with high-risk clinical characteristics in the Chinese population [12]. In this study, we also collected rectal cancer patients, which was nearly half of the patients. These would supplement rectal cancer patients from the Chinese population and proved the prognosis value of the automated analysis of ploidy, nucleotyping, and stroma fraction in Stage II CRC patients. In addition, we constructed a new model of PSN panel by combining the Ploidy, Stroma, and Nucleotyping. In this model, patients with diploidy, low-stroma, and chromatin homogeneous tumors were classified as the low-risk group, patients with three high-risk factors (non-diploidy, high-stroma, or chromatin heterogeneous) were classified into the high-risk group. The high-risk group patients had the shortest five years survival outcomes. These results indicated that the PSN panel could predict the prognosis of stage II CRC patients. We sought to validate our PSN panel in a separate cohort of 188 patient samples obtained from BJCH. Unlike the CAMSCH cohort, the PSN panel was only able to statistically distinguish the PSN-high-risk group (HR=3.551 [95% CI =1.411-8.938]), but not the PSN-middle-risk group in DFS. In the OS analysis, the stratification showed similar trends but was not statistically reliable. The key difference is likely due to the difference in clinical-pathological features between the two cohorts. The cases included from the BJCH-cohort were clinical-high-risk stage II cancers patients with at least one clinicopathological-high-risk feature, whereas the CAMSCH cohort includes all stage II CRC patients. This data proves that additional clinical features likely augmented the prognostic value of the PSN panel.

The T stage and microsatellite instability status are important clinical prognosis factors in CRC therapy. In this study, the combination of PSN factors with the T stage or MSI status could change the prognosis stratify of stage II CRC patients. The OS and DFS of T3 patients with PSN-high-risk factors were similar with T4 patients, and the prognosis of MSS patients with PSN-low-risk factors was better than that of MSI-H patients. These may indicate that the PSN factors can predict the prognosis of stage CRC II patients better.

The benefit from postoperative adjuvant therapy in Stage II CRC was less than 5% [24, 25]. The adjuvant treatment decisions for stage II colorectal cancer are currently dependent on the clinicopathological risk factors, which include the T stage, mismatch repair status, histological grade, lymph nodes samplings, etc. However, even high-risk patients grouped according to these risk factors cannot benefit

from postoperative adjuvant therapy or have limited benefits [3]. Therefore, the postoperative treatment of stage II patients is complex, and many researchers are also looking for biomarkers that can predict the efficacy of chemotherapy. In this study, postoperative adjuvant therapy did not improve OS or DFS in 179 stage II patients. When divided the patients according to the T3 or MSI status, the T3 patients who received adjuvant therapy showed a better OS (P=0.023), but the DFS was not significant in both T3 and MSS patients (**Table 3**). However, when patients were stratified using the PSN panel, postoperative adjuvant therapy significantly improved OS and DFS in patients with PSN-high-risk factors. In contrast, patients receiving postoperative adjuvant therapy in the PSN-low-risk group showed poorer survival, although this result was not statistically significant. According to these results, we infer that stratification of patients based on Ploidy, Stroma, and Nucleotyping would help select patients who may benefit from postoperative adjuvant therapy. In future studies, the sample size needs to be expanded to further confirm the role of PSN in guiding postoperative adjuvant therapy.

In the malignant tumor cells, the chromatin is less stable and the DNA is easier to access for the small molecular [26]. As shown in Figures S7 and S8, the adjuvant would significantly improve the OS in non-diploid (P=0.022) and chromatin heterogeneous (P=0.05) patients. The adjuvant chemotherapy drugs for stage II colorectal cancer are mainly fluorouracil and oxaliplatin, which target the DNA. Fluorouracil can inhibit the synthesis of DNA and RNA [27], while platinum drugs can inhibit the replication and transcription of DNA [28]. The chromatin is unstable in non-diploid or chromatin heterogeneous tumor cells; therefore, the small molecules and gamma-irradiation may exert anticancer activity easily. These may help us to understand the correlation between therapy efficacy and DNA status.

However this study did not find the role of stroma ratio in predicting the efficacy of adjuvant therapy. Recently, the function of the tumor stroma in chemotherapeutic resistance has attracted attention. Over the past decade, studies have profound the importance of the tumor microenvironment (TME) as a driving force in tumor development and progression. Not only does the desmoplastic stroma create a protective shield from therapeutics [29], it also aids in the process of epithelial to mesenchymal transition (EMT) and causes tumor cell dissemination into surrounding tissue. A recent study showed that the patients will get a better prognosis if the stroma fraction was deceased by neoadjuvant therapy [30]. The association between stroma ratio and therapy efficacy needs to be investigated further.

In a previous study [5], DNA ploidy and stroma stratification did not indicate the differences in the cancer-specific survival of Stage III CRC patients (HR=1.43, P=0.14). Besides, the stage III patients will receive adjuvant therapy after surgery according to the clinical guidelines, and the relationship between Ploidy, Stroma, Nucleotyping and adjuvant therapy was not clear. We were not sure about the possible value of PSN in stage III patients, so this study did not enroll stage III patients.

However, according to this study, stage II patients in the group of PSN-high risk group may benefit from adjuvant therapy and this provided us with some clues to explore the predictive value of PSN for adjuvant therapy.

Conclusion

Our study demonstrated that non-diploid, high-stroma, and chromatin heterogeneous were poor prognostic factors for stage II colorectal cancer patients. The ploidy, stroma-tumor-fraction, and nucleotyping as standalone parameters or combined panel are useful prognostic tools that can predict patient survival. In addition, patients with high-risk factors can benefit from adjuvant therapy, while patients without high-risk factors cannot. In future studies, the PSN panel can be further evaluated as a predictive biomarker to guide chemotherapy decisions. Lastly, the methods used in these studies are low cost and fairly easy to perform. Based on these factors, we believe that the PSN panel has the potential to be integrated into a routine pathology examination to aid in clinical-decision making for treating stage II CRC patients.

Acknowledgements

The authors would like to thank the donors and the Cancer Institute & Hospital, Chinese Academy of Medical Sciences, Peking Union

Medical College, Beijing, China for tumor specimens which was provided for the study, and acknowledge the efforts of Beijing Cancer Hospital, Beijing, China to provide the related data and information as validation cohort. The interpretation of these data is the sole responsibility of the authors. We also thank Professor David Kerr for the revision of the manuscript. We also would like to thank Doctor Lijun Mao, Chan Robin Barry, Fei Wang and Yanzheng Tong in Ningbo Meishan FTZ MBM Clinical Lab Co., Ltd. A. for providing technical support in ploidy, stroma, and Nucleotyping analysis. This study was supported by National Key R&D Program of China (2020AAA0109500), National Natural Science Foundation of China (82122053), the CAMS Innovation Fund for Medical Sciences (CIFMS) [2016-I2M-1-001, 2018-I2M-AI-008], Beijing Science and Technology Program [D17-110002617002, D17110002617004], and Beijing Hope Run Special Fund of Cancer Foundation of China [LC2017A07, LC2019B14].

Disclosure of conflict of interest

None.

Address correspondence to: Xishan Wang, Department of Colorectal Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 9, Dongdan San 3, Dongcheng District, Beijing, China. Tel: +86-17601099616; E-mail: wxshan1208@126.com; Shuangmei Zou, Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 9, Dongdan San 3, Dongcheng District, Beijing, China. E-mail: zousm@cicams.ac.cn; Aiwen Wu, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Department of Colorectal Surgery, Peking University Cancer Hospital & Institute, 52 Fucheng Road, Haidian District, Beijing, China. E-mail: wuaw@sina.com

References

[1] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. CA Cancer J Clin 2016; 66: 115-32.

[2] Weitz J, Koch M, Debus J, Höhler T, Galle PR and Büchler MW. Colorectal cancer. Lancet 2005; 365: 153-65.

[3] Kumar A, Kennecke HF, Renouf DJ, Lim HJ, Gill S, Woods R, Speers C and Cheung WY. Adjuvant chemotherapy use and outcomes of patients with high-risk versus low-risk stage II colon cancer. Cancer 2015; 121: 527-34.

[4] Danielsen HE, Pradhan M and Novelli M. Revisiting tumour aneuploidy-the place of ploidy assessment in the molecular era. Nat Rev Clin Oncol 2016; 13: 291-304.

[5] Danielsen HE, Hveem TS, Domingo E, Pradhan M, Kleppe A, Syvertsen RA, Kostolomov I, Nesheim JA, Askautrud HA, Nesbakken A, Lothe RA, Svindland A, Shepherd N, Novelli M, Johnstone E, Tomlinson I, Kerr R and Kerr DJ. Prognostic markers for colorectal cancer: estimating ploidy and stroma. Ann Oncol 2018; 29: 616-623.

[6] Van Pelt GW, Sandberg TP, Morreau H, Gelderblom H, van Krieken JHJM, Tollenaar RAEM and Mesker WE. The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. Histopathology 2018; 73: 197-206.

[7] Park JH, Richards CH, McMillan DC, Horgan PG and Roxburgh CSD. The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. Ann Oncol 2014; 25: 644-651.

[8] Huijbers A, Tollenaar RA, v Pelt GW, Zeestraten EC, Dutton S, McConkey CC, Domingo E, Smit VT, Midgley R, Warren BF, Johnstone EC, Kerr DJ and Mesker WE. The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. Ann Oncol 2013; 24: 179-85.

[9] Schuster-Böckler B and Lehner B. Chromatin organization is a major influence on regional mutation rates in human cancer cells. Nature 2012; 488: 504-7.

[10] Polak P, Karlić R, Koren A, Thurman R, Sandstrom R, Lawrence M, Reynolds A, Rynes E, Vlahoviček K, Stamatoyannopoulos JA and Sunyaev SR. Cell-of-origin chromatin organization shapes the mutational landscape of cancer. Nature 2015; 518: 360-364.

[11] Kleppe A, Albregtsen F, Vlatkovic L, Pradhan M, Nielsen B, Hveem TS, Askautrud HA, Kristensen GB, Nesbakken A, Trovik J, Wæhre H, Tomlinson I, Shepherd NA, Novelli M, Kerr DJ and Danielsen HE. Chromatin organisation and cancer prognosis: a pan-cancer study. Lancet Oncol 2018; 19: 356-369.

[12] Yang L, Chen P, Zhang L, Wang L, Sun T, Zhou L, Li Z and Wu A. Prognostic value of nucleotyping, DNA ploidy and stroma in high-risk stage II colon cancer. Br J Cancer 2020; 123: 973-981.

[13] Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL

and Siegel RL. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin 2019; 69: 363-385.
[14] Ji L and Tucker J. DNA measurement of overlapping cell nuclei in thick tissue sections. Anal Cell Pathol 1997; 14: 41-9.
[15] Pradhan M, Abeler VM, Danielsen HE, Tropé CG and Risberg BA. Image cytometry DNA ploidy correlates with histological subtypes in endometrial carcinomas. Mod Pathol 2006; 19: 1227-35.
[16] Haroske G, Baak JP, Danielsen H, Giroud F, Gschwendtner A, Oberholzer M, Reith A, Spieler P and Böcking A. Fourth updated ESACP consensus report on diagnostic DNA image cytometry. Anal Cell Pathol 2001; 23: 89-95.
[17] National Comprehensive Cancer Network. (NCCN) Clinical Practice Guidelines in Oncology. Colon Cancer, Version 4. 2020. https://www.nccn.org/professionals/physician_gls/pdf/colon.pdf. Accessed 15 June 2020.
[18] Kristensen GB, Kildal W, Abeler VM, Kaern J, Vergote I, Tropé CG and Danielsen HE. Large-scale genomic instability predicts long-term outcome for women with invasive stage I ovarian cancer. Ann Oncol 2003; 14: 1494-500.
[19] Pretorius ME, Waehre H, Abeler VM, Davidson B, Vlatkovic L, Lothe RA, Giercksky KE and Danielsen HE. Large scale genomic instability as an additive prognostic marker in early prostate cancer. Cell Oncol 2009; 31: 251-9.
[20] Hveem TS, Merok MA, Pretorius ME, Novelli M, Bævre MS, Sjo OH, Clinch N, Liestøl K, Svindland A, Lothe RA, Nesbakken A and Danielsen HE. Prognostic impact of genomic instability in colorectal cancer. Br J Cancer 2014; 110: 2159-64.
[21] Vermeulen L, De Sousa E Melo F, van der Heijden M, Cameron K, de Jong JH, Borovski T, Tuynman JB, Todaro M, Merz C, Rodermond H, Sprick MR, Kemper K, Richel DJ, Stassi G and Medema JP. Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. Nat Cell Biol 2010; 12: 468-76.
[22] Damato A, Iachetta F, Antonuzzo L, Nasti G, Bergamo F, Bordonaro R, Maiello E, Zaniboni A, Tonini G, Romagnani A, Berselli A, Normanno N and Pinto C. Phase II study on first-line treatment of NIVolumab in combination with folfoxiri/bevacizumab in patients with Advanced COloRectal cancer RAS or BRAF mutated-NIVACOR trial (GOIRC-03-2018). BMC Cancer 2020; 20: 822.
[23] Davoli T, Uno H, Wooten EC and Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. Science 2017; 355: eaaf83-99.
[24] Böckelman C, Engelmann BE, Kaprio T, Hansen TF and Glimelius B. Risk of recurrence in patients with colon cancer stage II and III: a systematic review and meta-analysis of recent literature. Acta Oncol 2015; 54: 5-16.
[25] Quasar Collaborative Group, Gray R, Barnwell J, McConkey C, Hills RK, Williams NS and Kerr DJ. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. Lancet 2007; 370: 2020-9.
[26] Gurova KV. Chromatin stability as a target for cancer treatment. Bioessays 2019; 41: e1800141.
[27] Hoskins J and Scott Butler J. Evidence for distinct DNA- and RNA-based mechanisms of 5-fluorouracil cytotoxicity in Saccharomyces cerevisiae. Yeast 2007; 24: 861-70.
[28] Slyskova J, Sabatella M, Ribeiro-Silva C, Stok C, Theil AF, Vermeulen W and Lans H. Base and nucleotide excision repair facilitate resolution of platinum drugs-induced transcription blockage. Nucleic Acids Res 2018; 12; 46: 9537-9549.
[29] Dauer P, Nomura A, Saluja A and Banerjee S. Microenvironment in determining chemo-resistance in pancreatic cancer: neighborhood matters. Pancreatology 2017; 17: 7-12.
[30] Hansen TF, Kjær-Frifeldt S, Lindebjerg J, Rafaelsen SR, Jensen LH, Jakobsen A and Sørensen FB. Tumor-stroma ratio predicts recurrence in patients with colon cancer treated with neoadjuvant chemotherapy. Acta Oncol 2018; 57: 528-533.

Supplementary Information

The inclusion criteria of the BJCH cohort

188 cases that had been diagnosed as high-risk stage II colon cancer in the Peking University Cancer Hospital from 2009 to 2015 were retrospectively enrolled. Patients in this cohort were with at least one of the following clinical pathology high risk features: lymph nodes sampling less than 12, poorly differentiated tumor, vascular or perineural invasion, pathological T4 stage, and clinical presentation with intestinal occlusion or perforation [10]. None of the patients received neoadjuvant chemotherapy or radiotherapy. Some patients received conventional adjuvant chemotherapy after surgery. The baseline comparison between BJCH-cohort and CAMSCH stage II cohort is summarized in the following Table.

Table S1. Univariate analysis of potential prognostic factors on 5-years overall survival and disease-free survival in the stage II CRC cohort

Factor	N	OS			DFS		
		5y OS (%)	HR (95% CI)	*P* value	5y DFS (%)	HR (95% CI)	*P* value
Age (years)				0.483			0.573
<65	93	75.3	1		68.8	1	
≥65	86	68.6	1.227 (0.639-2.172)		64.0	1.163 (0.688-1.966)	
Gender				0.149			0.55
Male	109	67.9	1		65.1	1	
Female	70	78.6	0.641 (0.350-1.173)		68.6	0.852 (0.504-1.44)	
Tumor location				0.252			0.681
Rectum	88	70.5	1		60.2	1	
Colon	91	73.6	1.518 (0.744-3.099)		72.5	1.414 (0.609-2.138)	
Gross pathological type				0.685			0.592
Prominence	105	70.5	1		64.8	1	
Ulceration & infiltration	74	74.3	0.889 (0.502-1.573)		68.9	0.867 (0.515-1.460)	
T stage				0.845			0.579
T3	156	72.4	1		67.3	1	
T4	23	69.6	1.083 (0.487-2.407)		60.9	1.222 (0.601-2.484)	
Histological grade				0.766			0.567
High	21	66.7	1		57.1	1	
Middle	149	73.2	0.766 (0.343-1.711)	0.516	67.8	0.667 (0.327-1.359)	
Low	9	66.7	0.998 (0.258-3.858)	0.997	66.7	0.682 (0.185-2.518)	
Adjuvant treatment				0.0598			0.187
Yes	112	76.8	1		69.6	1	
No	67	64.2	1.704 (0.978-2.969)		61.2	1.41 (0.846-2.351)	
MSI status				0.833			0.833
MSS/MSI-L	163	71.2	1		66.3	1	
MSI-H	16	81.3	0.906 (0.363-2.263)		68.8	0.906 (0.363-2.263)	
Ploidy				0.022			0.012
Diploidy	73	80.8	1		76.7	1	
Non-diploidy	106	66.0	2.057 (1.109-3.816)		59.4	2.063 (1.176-3.618)	
Stroma				0.006			0.082
Low stroma fraction	136	77.2	1		69.9	1	
High stroma fraction	43	55.8	2.244 (1.267-3.976)		55.8	1.621 (0.940-2.794)	
Nucleotyping				0.036			0.003
Chromatin homogeneous	125	76.8	1		73.6	1	
Chromatin heterogeneous	54	61.1	1.821 (1.038-3.194)		50.0	2.163 (1.300-3.601)	
Ploidy-Stroma				0.0019			0.0046
Diploid and low stroma (PS-L)	60	86.7	1		81.7	1	
Non-diploid or high stroma (PS-M)	89	67.4	2.808 (1.283-6.144)	0.0096	59.6	2.594 (1.320-5.097)	0.0057
Non-diploid and high stroma (PS-H)	30	56.7	4.228 (1.751-10.209)	0.0014	56.7	2.984 (1.335-6.667)	0.0077
Ploidy-Nucleotyping				0.0044			0.0009
Diploid and chromatin homogeneous (PN-L)	68	85.3	1		82.4	1	
Non-diploid or chromatin heterogeneous (PN-M)	62	62.9	2.994 (1.424-6.294)	0.0038	58.1	2.91 (1.468-5.772)	0.0022
Non-diploid and chromatin heterogeneous (PN-H)	49	65.3	2.752 (1.260-6.011)	0.0111	55.1	3.103 (1.534-6.274)	0.0016
Stroma-Nucleotyping				0.0035			0.0025
Low stroma and chromatin homogeneous (SN-L)	96	82.3	1		78.1	1	
High stroma or chromatin heterogeneous (SN-M)	69	62.3	2.415 (1.310-4.452)	0.0047	53.6	2.415 (1.392-4.192)	0.0017
High stroma and chromatin heterogeneous (SN-H)	14	50.0	3.407 (1.412-8.223)	0.0064	50.0	2.809 (1.193-6.618)	0.0181
Ploidy-Stroma-Nucleotyping				0.0005			0.0002
PSN-L	57	89.5	1		86.0	1	
PSN-M	53	67.9	3.547 (1.398-9.000)	0.0077	62.3	3.184 (1.402-7.232)	0.0057
PSN-H	69	60.9	4.554 (1.879-11.040)	0.0008	53.6	4.147 (1.909-9.007)	0.0003

PSN-L: Diploid and low stroma and chromatin homogeneous; PSN-M: Including one or two factors of Non-diploid, high stroma and heterogeneous; PSN-H: Non-diploid, high stroma and heterogeneous.

Predictive value of ploidy, stroma, and nucleotyping in stage II CRC

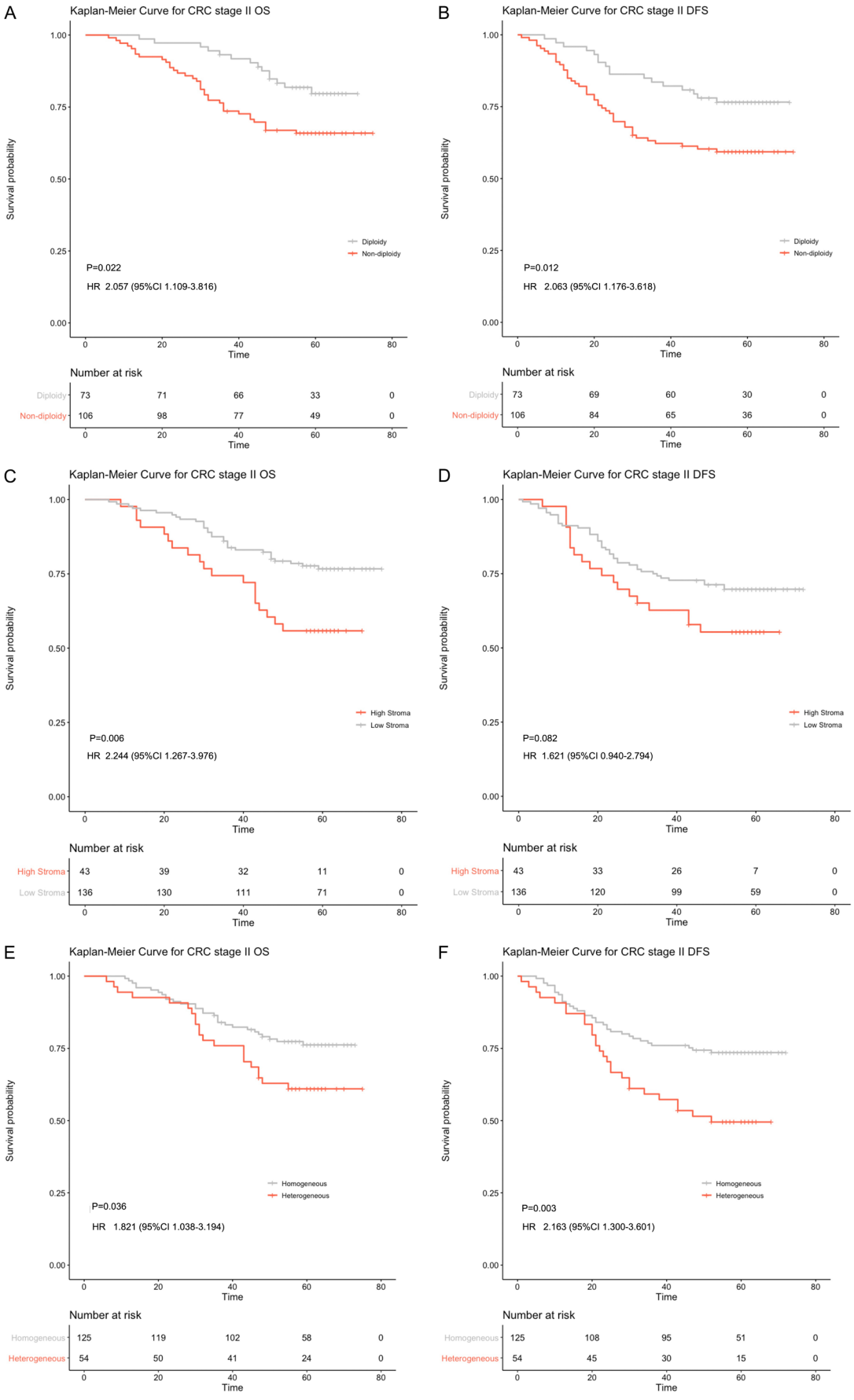
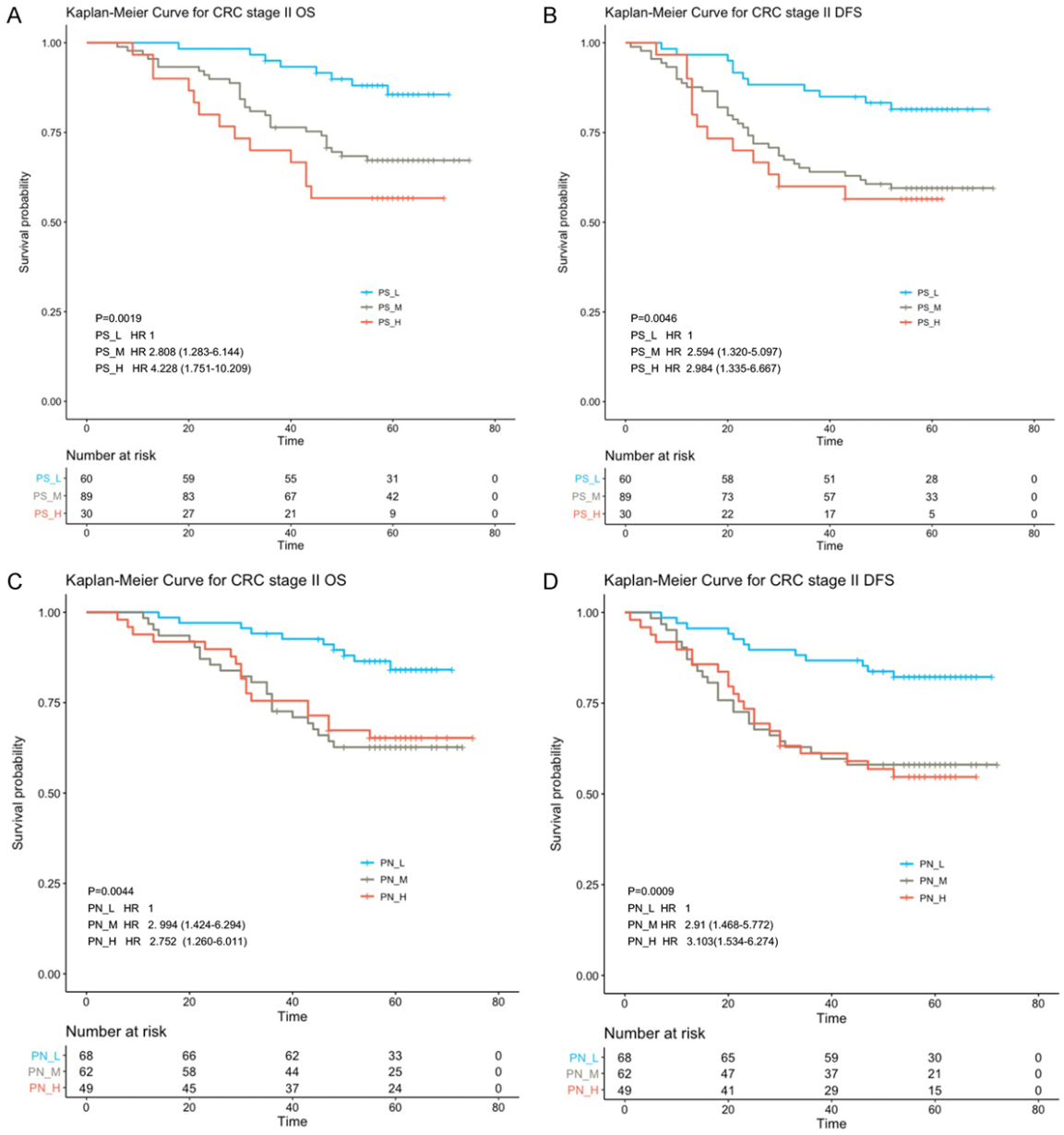


Figure S1. Kaplan-Meier plots of OS and DFS in stage II CRC patients based on the ploidy (A, B), stroma-tumor fraction (C, D), and chromatin organization (E, F).

Table S2. The correlation analysis between ploidy and nucleotyping

Variable	Chromatin homogeneous	Chromatin heterogeneous	Total	Coefficient	*P*-value
				0.336	<0.001
Diploid	150	7	157		
Non-diploid	207	132	339		
Total	357	139			



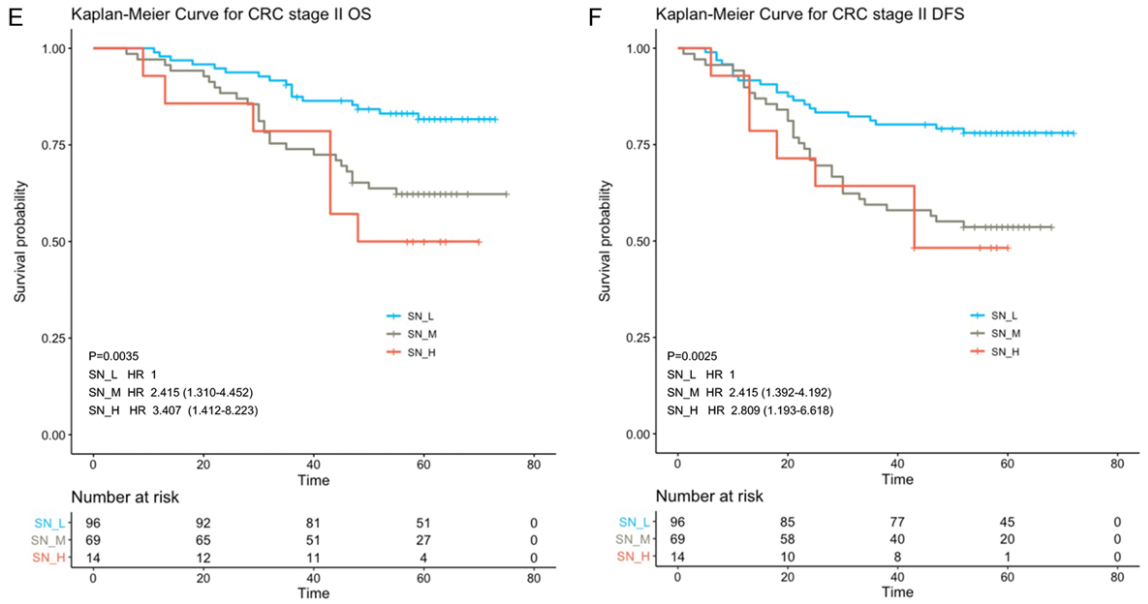


Figure S2. Kaplan-Meier plots illustrating OS and DFS for stage II CRC patients using different combinations of ploidy, stroma-tumor fraction and Nucleotyping. PS: the combination of ploidy and stroma-tumor fraction (A, B); PN: the combination of ploidy and Nucleotyping (C, D); SN: the combination of stroma-tumor fraction and Nucleotyping (E, F); L: low risk; H: high risk; M: middle risk.

Table S3. Comparison of demographic information between stage II CRC CAMSCH cohort and BJCH cohort

Factor	CAMSCH cohort stage II	BJCH cohort	Statistics	*P* value
Age (years)			0.323	0.570
≤63	91	90		
>63	88	98		
Gender			0.416	0.519
Male	110	121		
Female	70	67		
Grade			42.460	<0.001
High	21	7		
Middle	149	123		
Low	9	55		
T stage			21.820	<0.001
T3	156	125		
T4	23	63		
Adjuvant therapy			1.219	0.270
YES	112	107		
No	67	81		
MSI status			7.178	0.0074
MSS/MSS-L	163	153		
MSI-H	16	35		
Ploidy			1.352	0.245
Diploid	73	88		
Aneuploid + Tetraploid	106	100		
Tumor-stroma fraction			1.601	0.206
>0.5	43	35		

≤0.5	136	153		
Nucleotyping			6.015	0.014
≥0.044	125	152		
<0.044	54	36		

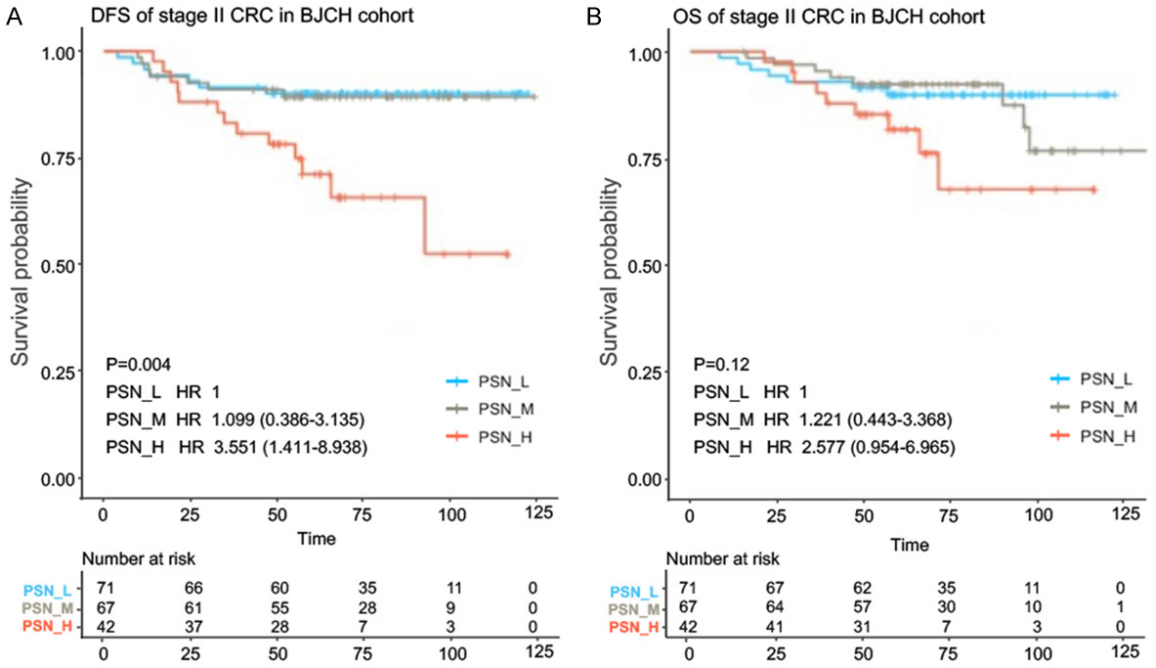


Figure S3. Kaplan-Meier plots illustrating DFS (A) and OS (B) of BJCH cohort when plotted according to the PSN panel classifier. L: low risk; H: high risk; M: middle risk.

Table S4. Kaplan-Meier analysis of the stage II CRC patients when stratified by the combination of Ploidy-Stroma-Nucleotyping (low-risk or middle-risk/high-risk)_with T staging or microsatellite instability status

Variables	Total	OS			DFS		
		Events	5y OS(%)	*P*-value	Events	5y DFS (%)	*P*-value
T&PSN				0.015			0.013
T3&PSN-LR	49	6	86.3		49	83.5	
T3&PSN-MR&HR	107	37	65.3		107	59.8	
T4	23	7	69.6		23	60.6	
MSI&PSN				0.005			0.003
MSS/MSI-L&PSN-LR	49	6	86.4		7	85.5	
MSS/MSI-L&PSN-MR&HR	114	41	63.9		48	57.8	
MSI-H	16	3	81.3		5	68.8	

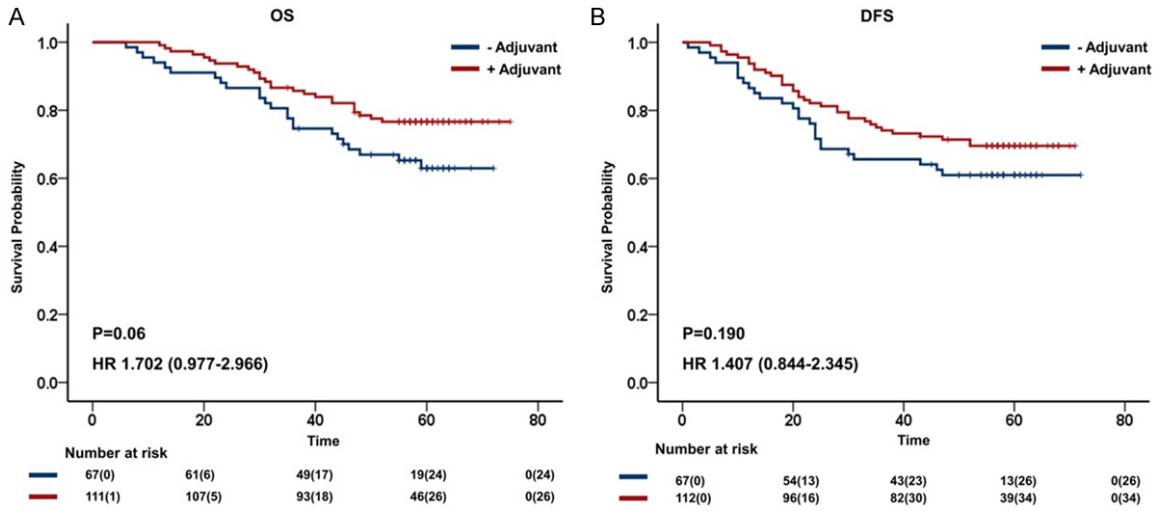


Figure S4. Kaplan-Meier plots illustrating OS (A) and DFS (B) of stage II patients who received adjuvant therapy versus those who did not.

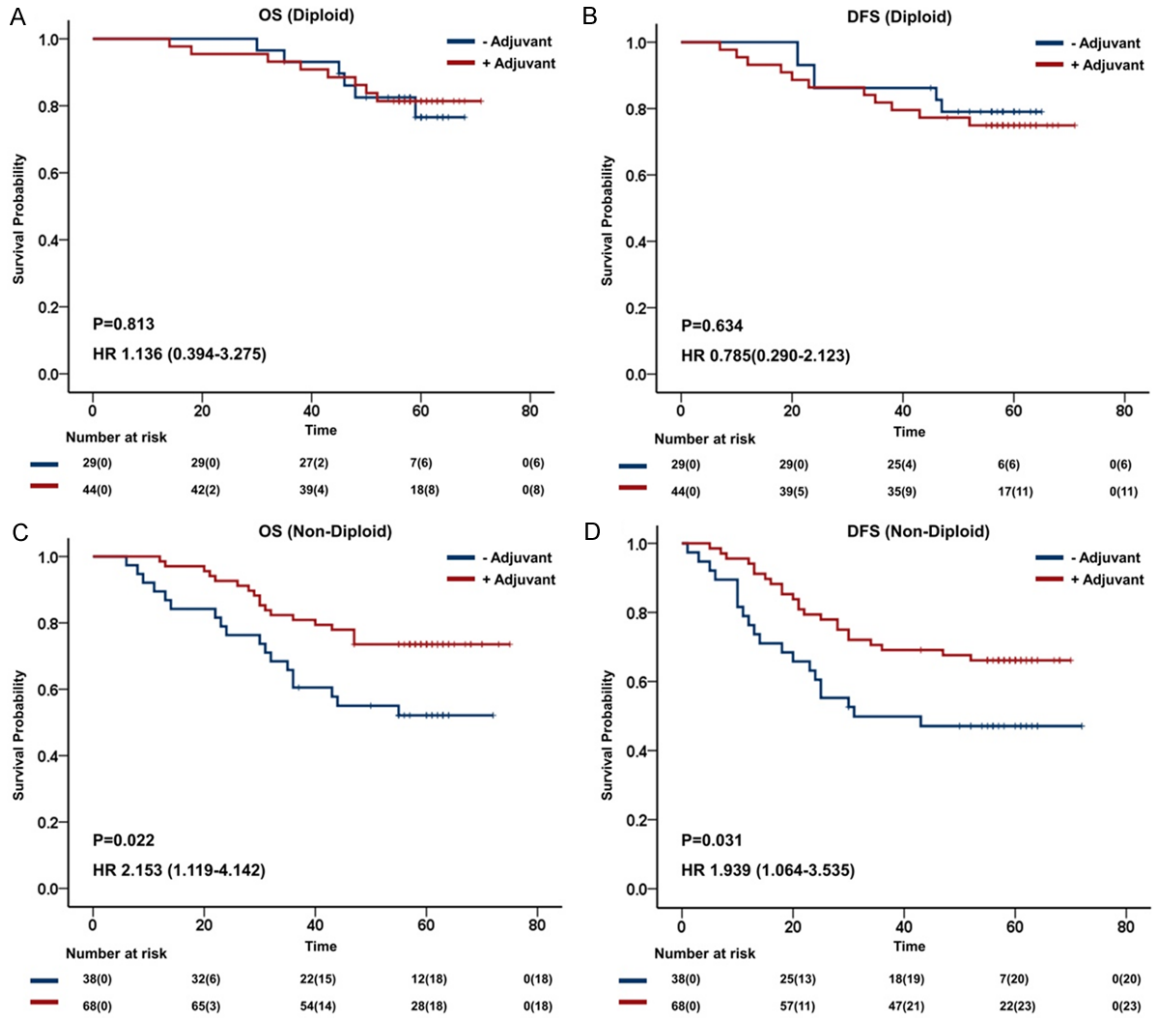


Figure S5. Kaplan-Meier plots illustrating the impact of adjuvant therapy on OS and DFS in stage II patients with tumors that were diploid (A, B) or non-diploid (C, D).

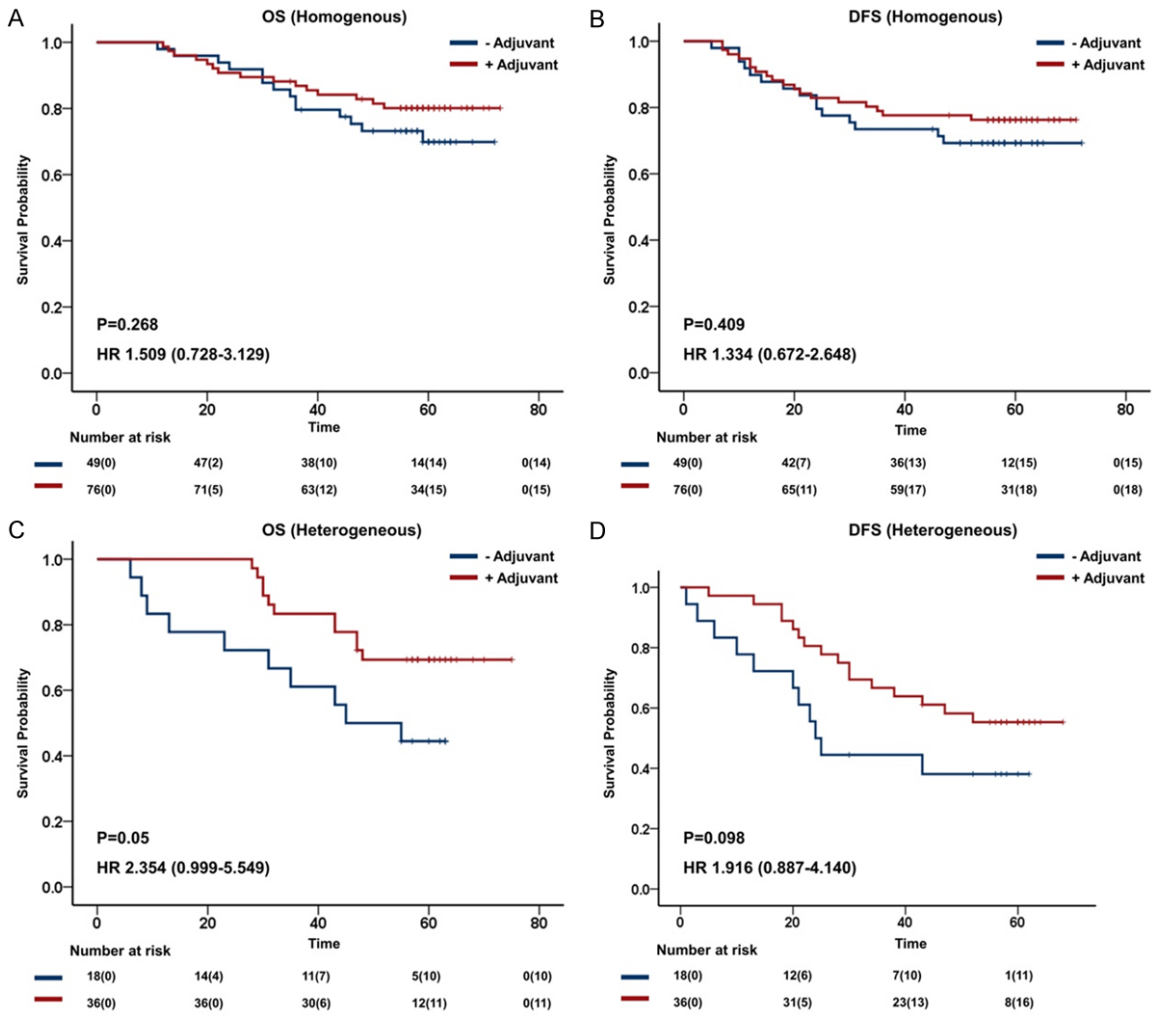


Figure S6. Kaplan-Meier plots illustrating the impact of adjuvant therapy on OS and DFS in stage II patients, with tumors that were chromatin homogeneous (A, B) or chromatin heterogeneous (C, D).

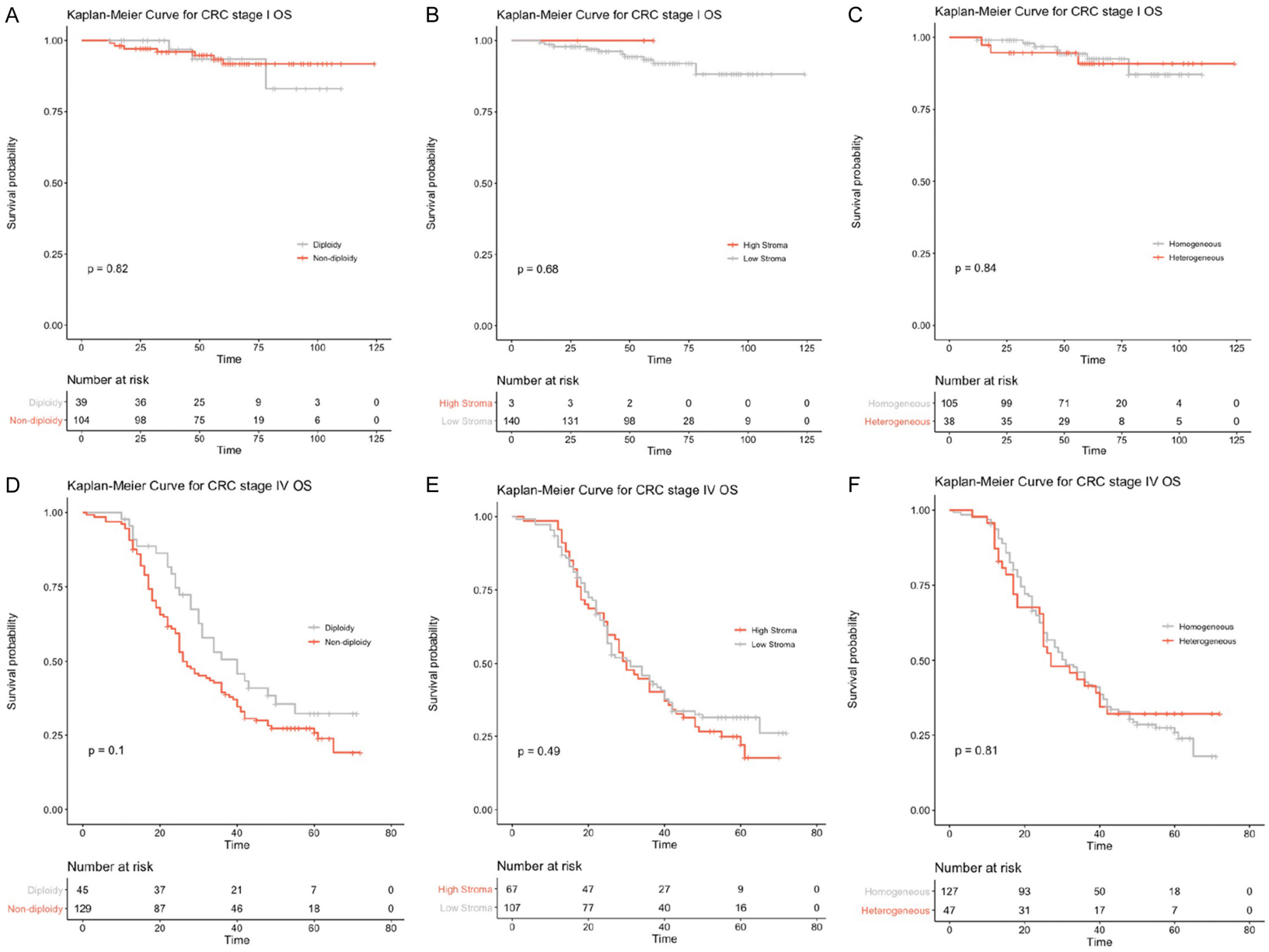


Figure S7. Kaplan-Meier plots illustrating overall survival of stage I (A-C) or liver metastatic CRC (D-F) patients, based on ploidy, stroma-tumor fraction, and chromatin organization.

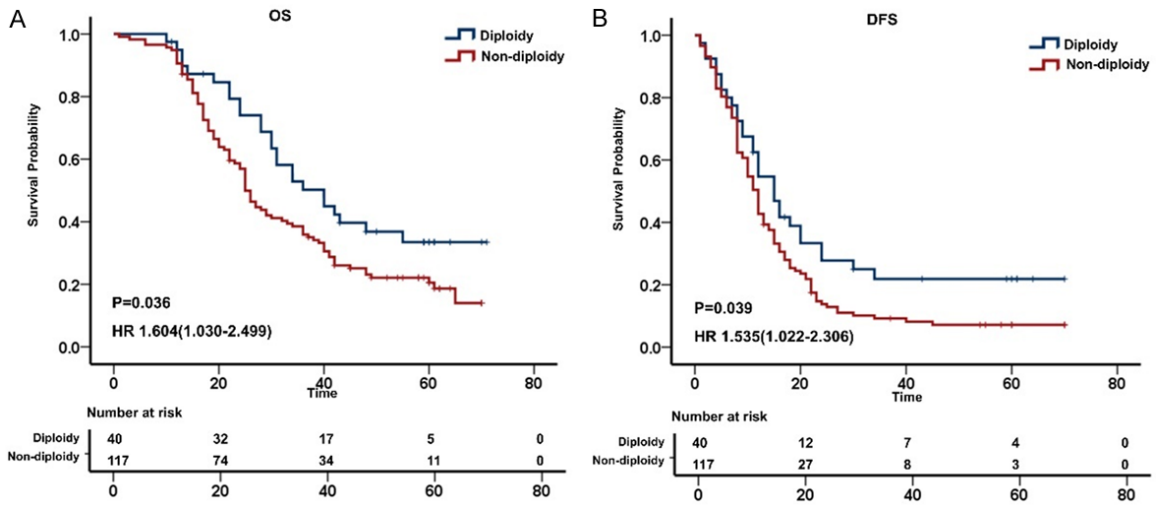


Figure S8. Kaplan-Meier plots illustrating OS (A) and DFS (B) for LMCRC patients with tumors that were microsatellite stable (MSS/MSI-L) and either diploid or non-diploid.