Human Pathology (2020) **105**, 37–46

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Unique clinicopathologic and genetic alteration features in early onset colorectal carcinoma compared with age-related colorectal carcinoma: a large cohort next generation sequence analysis☆,☆☆

David Escobar MD, PhD, Ryan Jones MD, PhD, Juehua Gao MD, PhD, Leyu Sun MD, Jie Liao MD, Guang-Yu Yang MD, PhD*

Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60611, USA

Received 9 June 2020; revised 4 August 2020; accepted 6 August 2020

Available online 9 September 2020

Keywords:
Early onset colorectal carcinoma;
Genetic alterations;
Clinicopathologic features;
PIK3CA;
KRAS;
CTNNB1;
NGS

Summary Colorectal carcinoma (CRC) is the third most common cancer type in the United States. While the incidence of CRC is decreasing among an older population undergoing screening, the incidence of early-onset CRC is rising. There is a growing understanding that the molecular underpinnings of colorectal carcinoma vary by age. In this study, we report the genetic alterations and clinicopathologic features of a single-institution colorectal carcinoma cohort over a 2-year period using a next-generation sequencing (NGS) approach and microsatellite stability (MS) status determined by immunohistochemical staining. Forty cases were identified in an early-onset colorectal carcinoma cohort (eCRC) defined by age <40 years, and 164 cases were identified in an age-related colorectal carcinoma cohort (arCRC) defined by age >70 years. eCRC was more often-left-sided/rectal and more likely to present high rates of lymph node positivity with metastatic disease. NGS mutational analysis revealed distinct differences between eCRC and arCRC, with eCRC being characterized by low frequency of *PIK3CA* mutations, elevated frequency of *KRAS* and *CTNNB1* mutations in microsatellite instability high tumors, and very low frequency of *BRAF* mutations.

☆ A portion of this work was presented as a poster (Abstract No. 1837) at the United States and Canadian Academy of Pathology (USCAP) 2020 Annual Meeting in Los Angeles, CA, USA during the Stowell-Orbison poster award session and was awarded a Stowell-Orbison Certificate of Merit.

☆☆ Competing interests: None

* Corresponding author. Northwestern University Feinberg School of Medicine, Department of Pathology, Ward 4-115 303 E Chicago Avenue, Chicago Illinois 60611, USA.

E-mail address: g-yang@northwestern.edu (G.-Y. Yang).

https://doi.org/10.1016/j.humpath.2020.08.002

1. Introduction

Colorectal carcinoma (CRC) is the third most common cancer in the United States [1]. In 2020, there will be an estimated 147 950 new diagnoses of colorectal cancer in the United States. Incidence rates are decreasing for individuals aged 50 years or older and overall death rates from CRC are decreasing, due in large part to screening regimens in industrialized nations. In contrast, CRC rates are increasing among younger individuals, especially in high-income countries [2]. The younger population may not be routinely screened for colorectal cancer and are at risk for more advanced disease at time of diagnosis [3,4], with worse progression-free survival and overall survival compared with older patients [5]. Reflective of the increasing incidence of CRC in a younger population, the median age of diagnosis has decreased to 66 years in men and 69 years in women [1].

While a subset of early-onset colorectal carcinoma (eCRC) is attributable to an identifiable predisposing genetic condition (ie, hereditary nonpolyposis colorectal cancer/Lynch Syndrome and familial adenomatous polyposis [FAP]) [6–8], hereditary conditions cannot fully account for the rise in incidence among the young. There is an association between chronic inflammatory bowel disease (IBD) and the development of carcinoma [9,10]; however, many eCRCs may arise sporadically [6,11].

While the genetic differences between early onset and age-related colorectal carcinoma (aCRC) continue to be an area of investigation, knowledge about the genetic alterations in eCRC remains confined to a few studies to date. With burgeoning interest in the molecular and genetic features of early-onset colorectal carcinoma, in this study, we report the genetic alterations and clinicopathologic features of a single-institution colorectal carcinoma cohort using a next-generation sequencing (NGS) approach, aiming to further elucidate if there are any unique genetic and morphologic features in eCRC.

2. Methods

2.1. Patient cohort

The pathology database at Northwestern Memorial Hospital in Chicago, IL, was searched between the years 2016 and 2018, and 206 individual patients (42 below the age of forty and 164 above the age of seventy) were identified with a diagnosis of colorectal cancer on either biopsy or resection with concurrent evaluation for mismatch repair (MMR) by immunohistochemistry (see "Microsatellite Instability Mismatch Repair Protein Expression" and somatic mutation using NGS using a 22-gene panel approach (see "Next-Generation Sequencing," below). Presence or absence of germline mutations, including *APC*, *MUTYH/NTLH1*, or *TP53,* was not used as either inclusion or exclusion criteria.

2.2. Microsatellite Instability Mismatch Repair Protein Expression

All tumors were evaluated for MMR deficiency and microsatellite instability (MSI) via immunohistochemical approach in a Clinical Laboratory Improvement Amendments–approved laboratory for clinical care. Immunohistochemical staining was performed using an automated immunostainer (Leica Bond-III; Leica Biosystems, Buffalo Grove, IL) and Bond Refine PolymerTM biotin-free 3,3′-diaminobenzidine detection kit. Antibodies included MLH1 (mouse anti-human antibody clone ES05, Leica, Buffalo Grove, IL), MSH2 (mouse anti-human antibody clone G219-1129, Cell-Marque, Rocklin, CA), MSH6 (mouse anti-human clone 44, Cell-Marque, Rocklin, CA), and PMS2 (mouse anti-human antibody clone MRQ-28, Cell-Marque, Rocklin, CA). Bright signal intensity in greater than 1% of tumor cells was considered positive for protein expression [12,13].

2.3. Next generation sequencing

Next generation sequencing was performed as described previously [14]. In brief, after review of the hematoxylin-eosin-stained slide from the same block was reviewed by a pathologist, DNA was extracted from FFPE tumor samples using the High Pure FFPE tissue kit (Roche, Indianapolis, IN) using 5-μm paraffin sections. The paraffin was removed with xylene, and the DNA was extracted according to the manufacturer's instructions. The DNA obtained was quantified using the Qubit DNA HS assay kit (Life Technologies, Carlsbad, CA).

For library construction, 10 ng of DNA (measured using the Qubit DNA HS assay kit) was amplified using the Lung and Colon Cancer panel V2 (Ampliseq; Life Technologies) and the Ion Ampliseq HiFi Master Mix (Ion Ampliseq Library kit 2.0) [15]. This panel of 1825 hotspot mutations in 22 genes was validated for colorectal carcinoma (n = 51) and nonsmall cell lung carcinoma (n = 39) as described in D'Haene et al. [16]. These commercially available products were then validated for use in solid tumors at the Diagnostic Molecular Biology laboratory at Northwestern Memorial Hospital. An amplicon library was thus generated for sequencing 1825 hotspot mutations in 22 genes, including AKT1 (NM_05163), ALK (NM_004304), BRAF (NM_004333), CTNNB1 (NM_001904), DDR2 (NM_001014796), EGFR (NM_005228), ERBB2 (NM_004448), ERBB4 (NM_005235), FBXW7 (NM_033632), FGFR1 (NM_023110), FGFR2 (NM_022970), FGFR3 (NM_000142), KRAS (NM_033360), MAP2K1 (NM_002755), MET (NM_001127500), NOTCH1 (NM_017617), NRAS (NM_002524), PIK3CA (NM_006218), PTEN

(NM_000314), SMAD4 (NM_005359), STK11 (NM_000455), and TP53 (NM_000546). The amplicons were then digested, barcoded, and amplified using the Ion Ampliseq Library kit 2.0 and Ion Xpress barcode adapter's kit (Life Technologies) according to the manufacturer's instructions. The library prepared was quantified using the Qubit fluorometer and the Qubit dsDNA HS assay kit (Life Technologies). 8pM of each library was multiplexed and clonally amplified on the Ion Chef using Hi-Q chemistry (Life Technologies) according to the manufacturer's instructions. Finally, the template ISP was loaded on an Ion 316 or on an Ion 318 chip and sequenced on a PGM sequencer with the Ion PGM Hi-Q sequencing according to the manufacturer's instructions.

The raw data were analyzed using the Torrent Suite software v5.2.1 (Life Technologies). Human genome build 19 was used as the reference for alignment. Identification of sequence variants was facilitated via IT Variant Caller Plugin software v5.2.1. (Life Technologies), and coverage of each amplicon was obtained by the Coverage Analysis Plugin software v5.2.1 (Life Technologies). Cases for which the number of mapped reads was less than 100 000 and/or the average base coverage was less than 300X were considered as noninformative.

The Integrative Genomics Viewer from the Broad Institute (http://www.broadinstitute.org/igv/) was used to visualize the read alignment and the presence of variants against the reference genome as well as to confirm variant calls by checking for strand biases and sequencing errors [17,18]. Only mutations reported in the Sanger Institute Catalogue of Somatic Mutations in Cancer database (http://www.sanger.ac.uk/cosmic) were taken into account, and silent or intronic mutations were not reported [19].

The use of paraffin blocks for this study meets institutional review board and Health Insurance Portability and Accountability Act requirements and has been approved by the institutional review board at the Northwestern University Feinberg School of Medicine.

2.4. Clinical data

Clinical parameters, including association with IBD, chemotherapy and radiation therapy data, and outcomes data were collected and managed using a collaborative REDCap™ (Vanderbilt University, Nashville TN, USA, funded in part by National Institutes of Health [NIH/NCATS UL1 TR000445]) electronic data capture tool hosted at the Northwestern University Clinical and Translational Sciences Institute (NUCATS) [20]. Clinical data were obtained and added to the REDCap database in a manner approved by the institutional review board at the Northwestern University Feinberg School of Medicine. Additionally; instances in which the REDCap database was incomplete in regard to the association with IBD, the pathology database was searched for prior colon biopsy specimens for each case. Any history of biopsy-proven IBD was reported as present. Any patient with at least one colon biopsy specimen that was negative for IBD at least 1 year before the date of diagnosis of carcinoma was reported as absent. All other cases were reported as IBD status unknown/not reported.

2.5. Statistical analysis

All calculations were carried out in Microsoft Excel (Microsoft Corporation, Redmond, WA) or GraphPad Prism (GraphPad, San Diego, CA). Frequencies, percentiles, median, range, Student t-test, Chi-square, and Fischer exact tests were utilized, as appropriate. Statistical significance was defined as $P < 0.05$.

3. Results

3.1. Pathological characteristics of the early-onset and age-related cohorts of CRC

Patient demographics and baseline characteristics are reported in Table 1. Two hundred six cases with MMR immunohistochemical staining status and NGS sequencing data were identified and sorted into an age-related colorectal carcinoma cohort (arCRC) defined as >70 years old (n = 164) and an early-onset colorectal carcinoma cohort (eCRC) defined as <40 years old (n = 42). Compatible with prior reports, eCRC was more often left-sided or rectal in location, whereas arCRC was more often right-sided (overall $P < 0.01$, Table 1). Early CRC was more likely to have advanced disease with higher rates of positive regional lymph nodes (overall $P < 0.001$) and metastatic disease with higher overall clinical tumor stage at time of diagnosis (overall $P < 0.05$, Table 1). Presence of prior IBD, a known risk factor for the development of colorectal carcinoma, was known to be present in three cases (7%) of the eCRC cohort, whereas IBD was known to be present in only one (0.6%) of the arCRC cohort. Overall, CRC in both the early-onset and age-related cohorts was predominantly adenocarcinoma of no special type (Table 1); however, there was a significantly higher proportion of adenocarcinomas of micropapillary subtype in the early CRC cohort ($P < 0.005$). We did not observe a significant difference in tumor size and overall histologic grade (Table 1) when comparing eCRC and arCRC.

3.2. Analysis on mismatch repair protein deficiency

To report the microsatellite stability status of CRC, the immunohistochemical expression of MMR proteins MLH1, MSH2, MSH6, and PMS2 was reported as microsatellite stable (MSS) or microsatellite instability high (MSI-H) (*see Methods*). Tumors were characterized as MLH1-deficient if negative for both MLH1 and PMS2 and as MSH2 deficient if negative for both MSH2 and MSH6. The overall rates of MSI-H tumors were similar between the eCRC and arCRC

Table 1 Overall pathological characteristic of the early-onset and age-related cohorts.

Case Parameters	Total Cases			MSS			MSI-H		
	<40 years	> 70 years	*P*	<40 years	>70 years	*P*	< 40 years	> 70 years	*P*
Cases	42	164	–	35	134		7	30	
Age (years)	35 (19–40)	77 (70–95)	–	34 (19–39)	77 (70–95)	–	36 (31–40)	77 (70–87)	–
Sex									
Male	21 (50%)	83 (50.6%)	–	17 (48.6%)	75 (56%)	0.452	4 (57.1%)	8 (27%)	0.183
Female	21 (50%)	81 (49.4%)	–	18 (51.4%)	59 (44%)	0.452	3 (42.9%)	22 (73%)	0.183
Location									
Right	7 (16.7%)	73 (44.5%)	**<0.001**	3 (8.6%)	50 (37.3%)	**<0.001**	4 (57.1%)	23 (76.7%)	0.360
Left	19 (45.2%)	47 (28.7%)	**0.044**	18 (51.4%)	45 (33.6%)	0.076	1 (14.3%)	2 (6.7%)	0.478
Transverse	2 (4.8%)	12 (7.3%)	0.740	1 (2.9%)	8 (6.0%)	0.687	1 (14.3%)	4 (13.3%)	>0.999
Rectum	13 (31.0%)	27 (16.5%)	**0.048**	12 (34.3%)	26 (19.4%)	0.071	1 (14.3%)	1 (3.3%)	0.347
Appendix	0 (0%)	1 (0.6%)	>0.999	0 (0%)	4 (3.0%)	0.582	0 (0%)	0 (0%)	–
Unlabeled	1 (2.4%)	4 (2.4%)	–	1 (2.9%)	1 (0.7%)	0.372	0 (0%)	0 (0%)	–
		Overall *P*	**0.009**		Overall *P*	**0.008**		Overall *P*	0.579
Tumor size (cm)	4.6 (0.2 −11.5)	4.7 (0.2 −14.5)	0.912	4.9 (0.8 −11.5)	4.6 (0.2 −12)	0.653	3.7 (02. - 6.5)	5.1 (1.1 −14.5)	0.335
Histologic grade									
Low/Mod	32 (76.2%)	113 (68.9%)	0.450	26 (74.3%)	101 (75.4%)	>0.999	**6 (85.7%)**	**12 (40.0%)**	**0.042**
High	10 (23.8%)	38 (23.2%)	>0.999	9 (25.7%)	22 (16.4%)	0.224	1 (14.3%)	16 (53.3%)	0.098
Not resected/Not reported	0 (0%)	13 (7.9%)	–	0 (0%)	11 (8.2%)	–	0 (0%)	2 (6.7%)	–
		Overall *P*	0.166		Overall *P*	0.123		Overall *P*	0.091
pN									
pNX	14 (33.3%)	40 (24.4%)	0.244	12 (34.3%)	35 (26.1%)	0.397	3 (42.9%)	5 (16.7%)	0.156
pN0	10 (23.8%)	88 (53.7)	**0.001**	8 (22.9%)	70 (52.2%)	**0.002**	2 (28.6%)	18 (60.0%)	0.212
pN1	7 (16.7%)	23 (14.0%)	0.631	7 (20.0%)	18 (13.4%)	0.421	0 (0%)	5 (16.7%)	0.560
pN2	11 (26.2%)	13 (7.9%)	**0.002**	8 (22.9%)	11 (8.2%)	**0.030**	2 (28.6%)	2 (6.7%)	0.155
		Overall *P*	**<0.001**		Overall *P*	**0.008**		Overall *P*	0.089
Metastases									
pMX/pM0	27 (64.3%)	128 (78.0.7%)	0.074	21 (60.0%)	100 (74.6%)	0.096	6 (85.7%)	29 (96.7%)	0.347
pM1	15 (35.7%)	36 (22.0%)	0.074	14 (40.0%)	34 (25.4%)	0.096	1 (14.3%)	1 (3.3%)	0.347
		Overall *P*	0.065		Overall *P*	0.088		Overall *P*	0.249
Tumor stage									
0	1 (2.4%)	4 (2.4%)	>0.999	0 (0%)	3 (2.2%)	>0.999	1 (14.3%)	1 (3.3%)	0.347
I	4 (9.5%)	31 (18.9%)	0.174	4 (11.4%)	26 (19.4%)	0.329	0 (0%)	5 (16.7%)	0.560
II	6 (14.3%)	51 (31.1%)	**0.033**	4 (11.4%)	39 (29.1%)	**0.048**	2 (28.6%)	12 (40.0%)	0.687
III	12 (28.6%)	26 (15.9%)	0.074	10 (28.6%)	20 (14.9%)	0.081	2 (28.6%)	6 (20.0%)	0.631
IV	15 (35.7%)	34 (20.7%)	0.066	14 (40.0%)	31 (23.1%)	0.054	1 (14.3%)	3 (10.0%)	0.488
Not Resected	4 (9.5%)	18 (11.0%)	–	3 (8.6%)	15 (11.2%)	>0.999	0 (0%)	3 (10.0%)	>0.999
		Overall *P*	**0.025**		Overall *P*	**0.042**		Overall *P*	0.570
Tumor subtype									
Adenocarcinoma NOS	33 (78.6%)	139 (84.8%)	0.354	28 (82.4%)	115 (85.8%)	0.432	4 (57.1%)	23 (76.7%)	0.360
Mucinous	1 (2.4%)	12 (7.3%)	0.305	0 (0%)	9 (6.7%)	0.207	1 (14.3%)	3 (10.0%)	>0.999
Micropapillary	6 (14.3%)	2 (1.2%)	**0.001**	4 (11.4%)	2 (1.5%)	**0.017**	2 (28.6%)	0 (0%)	**0.032**
Signet Ring	2 (4.8%)	10 (6.1%)	>0.999	2 (5.7%)	6 (4.5%)	0.671	0 (0%)	4 (13.3%)	0.570
Adenosquamous	–	1 (0.6%)	–	–	1 (0.7%)	–	–	–	–
Enteroblastic	–	1 (0.6%)	–	–	1 (0.7%)	–	–	–	–
		Overall *P*	**0.001**		Overall *P*	**0.016**		Overall *P*	**0.020**
IBD associated									
Yes	3 (7.0%)	1 (0.6%)	**0.027**	3 (8.6%)	1 (0.7%)	**0.028**	0 (0%)	0 (0%)	–
No/Unreported	39 (93.0%)	163 (99.4%)	**0.027**	32 (91.4%)	133 (99.3%)	**0.028**	7 (100%)	30 (100%)	–
		Overall *P*	**0.006**			**0.007**		Overall *P*	–

Note. Comparisons that reached statistical significance by Fischer's exact test or Chi-square test are bolded.
Abbreviation: MSS, microsatellite stable; MSI-H, microsatellite instability high; IBD, inflammatory bowel disease.

Table 2 Number of genes mutated and overall frequencies of genetic alterations detected in the early-onset and age-related cohorts, showing stratification by MSI-status.

	Total Cases			MSS			MSI-H		
	<40 years	> 70 years	*P*	<40 years	> 70 years	*P*	<40 years	> 70 years	*P*
Cases	42	164	–	35 (83.3%)	134 (81.7%)		7 (16.7%)	30 (18.3%)	
Num. genes mutated									
0	4 (9.5%)	15 (9.1%)	1.0000	4 (11.4%)	13 (9.7%)	0.756	0 (0%)	2 (6.7%)	>0.999
1	13 (31.0%)	43 (26.2%)	0.5625	12 (34.3%)	40 (29.9%)	0.682	1 (14.3%)	3 (10.0%)	>0.999
2	13 (31.0%)	61 (37.2%)	0.4778	10 (28.6%)	49 (36.6%)	0.430	3 (42.9%)	12 (40.0%)	>0.999
3+	12 (28.6%)	45 (27.4%)	>0.999	9 (25.7%)	32 (23.9%)	0.824	3 (42.9%)	13 (43.3%)	>0.999
					Overall *P*	0.734		Overall *P*	0.626
Genes mutated									
KRAS	19 (45.2%)	70 (42.6%)	0.862	14 (40.0%)	66 (49.3%)	0.349	5 (71.4%)	4 (13.3%)	**0.005**
EGFR	1 (2.4%)	4 (2.4%)	>0.999	0 (0%)	2 (1.5%)	>0.999	1 (14.3%)	2 (6.7%)	0.478
BRAF	2 (4.8%)	32 (19.5%)	**0.020**	2 (5.7%)	10 (7.5%)	>0.999	0 (0%)	21 (70.0%)	**0.001**
>*BRAF* V600E	1 (2.4%)	28 (17.1%)	**0.012**	1 (2.9%)	7 (5.2%)	0.347	0 (0%)	21 (70.0%)	**0.001**
PIK3CA	4 (9.5%)	48 (29.3%)	**0.009**	2 (5.7%)	34 (25.4%)	**0.010**	2 (28.6%)	14 (46.7%)	0.675
AKT1	1 (2.4%)	2 (1.2%)	0.497	1 (2.9%)	1 (0.7%)	0.372	0 (0%)	1 (3.3%)	>0.999
ERBB2	0 (0%)	3 (1.8%)	>0.999	0 (0%)	3 (2.2%)	>0.999	0 (0%)	0 (0%)	>0.999
PTEN	3 (7.1%)	9 (5.5%)	0.713	2 (5.7%)	4 (3.0%)	0.605	1 (14.3%)	5 (16.7%)	>0.999
NRAS	1 (2.4%)	4 (2.4%)	>0.999	1 (2.9%)	4 (3.0%)	>0.999	0 (0%)	0 (0%)	>0.999
STK11	0 (0%)	2 (1.2%)	>0.999	0 (0%)	1 (0.7%)	>0.999	0 (0%)	1 (3.3%)	>0.999
MAP2K1	0 (0%)	0 (0%)	–	–	–	–	–	–	–
ALK	0 (0%)	1 (0.6%)	>0.999	0 (0%)	1 (0.7%)	>0.999	0 (0%)	0 (0%)	>0.999
DDR2	1 (2.4%)	1 (0.6%)	0.373	1 (2.9%)	1 (0.7%)	>0.999	0 (0%)	1 (3.3%)	>0.999
CTNNB1	5 (11.9%)	4 (2.4%)	**0.019**	1 (2.9%)	3 (2.2%)	>0.999	4 (57.1%)	1 (3.3%)	**0.003**
MET	3 (7.1%)	7 (4.3%)	0.429	3 (8.6%)	5 (3.7%)	0.365	0 (0%)	2 (6.7%)	>0.999
TP53	27 (64.3%)	97 (59.1%)	0.599	26 (74.3%)	88 (65.7%)	0.419	1 (14.3%)	9 (30.0%)	0.647
SMAD4	4 (9.5%)	18 (11.0%)	>0.999	4 (11.4%)	13 (9.7%)	0.756	0 (0%)	5 (16.7%)	0.560
FBX7	2 (4.8%)	16 (9.8%)	0.539	2 (5.7%)	13 (9.7%)	0.739	0 (0%)	3 (10.0%)	>0.999
FGFR3	1 (2.4%)	2 (1.2%)	0.497	1 (2.9%)	2 (1.5%)	0.504	0 (0%)	0 (0%)	>0.999
NOTCH1	0 (0%)	2 (1.2%)	>0.999	0 (0%)	0 (0%)	>0.999	0 (0%)	2 (6.7%)	>0.999
ERBB4	2 (4.8%)	2 (1.2%)	0.185	1 (2.9%)	0 (0%)	>0.999	1 (14.3%)	2 (6.7%)	0.478
FGFR1	1 (2.4%)	0 (0%)	0.204	0 (0%)	0 (0%)	>0.999	1 (14.3%)	0 (0%)	0.189
FGFR2	0 (0%)	1 (0.6%)	>0.999	0 (0%)	0 (0%)	>0.999	0 (0%)	1 (3.3%)	>0.999
					Overall *P*	0.451		Overall *P*	**0.002**

Note. Comparisons that reached statistical significance by Fischer's exact test or Chi-square test are bolded.
Abbreviation: MSS, microsatellite stable; MSI-H, microsatellite instability high.

cohorts (16.7% versus 18.3%, Table 2). In eCRC, MSS tumors tended to be left-sided with higher rates of regional lymph node positivity; whereas in arCRC, MSS tumors were significantly more likely to be right-sided (Table 1). In contrast, in MSI-H tumors, the early and age-related cohorts showed no statistically significant difference in sidedness or regional lymph node positivity (Table 1). In MSI-H tumors, MLH1 deficiency was identified in 85.7% and 90.0% of eCRC and arCRCs, respectively, and MSH2 deficiency was identified in 14.3% and 10.0% of MSI-H tumors in eCRC and arCRC, respectively (not significantly different, Supplementary Table 1).

3.3. NGS mutational analysis

DNA was extracted from formalin-fixed, paraffin-embedded sections, and next-generation sequencing was performed using a 22 gene panel as described in the Methods. The distribution of genetic alterations identified is shown in Fig. 1A and Table 2 (complete table of annotated mutations are shared in Supplementary Tables 2 and 3). The number of genes mutated per tumor was similar when comparing the eCRC and arCRC cohorts (Fig. 1B). Overall, the most frequently mutated genes in both cohorts were *TP53*, *KRAS*, *PIK3CA*, *BRAF*, *SMAD4*, *FBX7*, *PTEN*, *MET*, and *CTNNB1* (Fig. 1A). The proportion of *PIK3CA* mutation was lower in eCRC compared with arCRC (9.5% versus 29.3%, $P = 0.009$, Table 2). Overall, the rate of *BRAF* mutations in the eCRC cohort was 4.8% compared with 19.5% in the arCRC cohort ($P = 0.020$, Table 2). Of these, *BRAF* V600E comprised 50% (1 out of 2) of the overall *BRAF* mutations in the eCRC cohort (2.4% overall). In contrast *BRAF* V600E comprised 87.5% (28 out of 32) of the overall *BRAF* mutations in the arCRC cohort (17.1% overall) ($P = 0.012$, Fig. 1A inset, Table 2). The rate of *CTNNB1* mutation was distinctly higher in eCRC compared

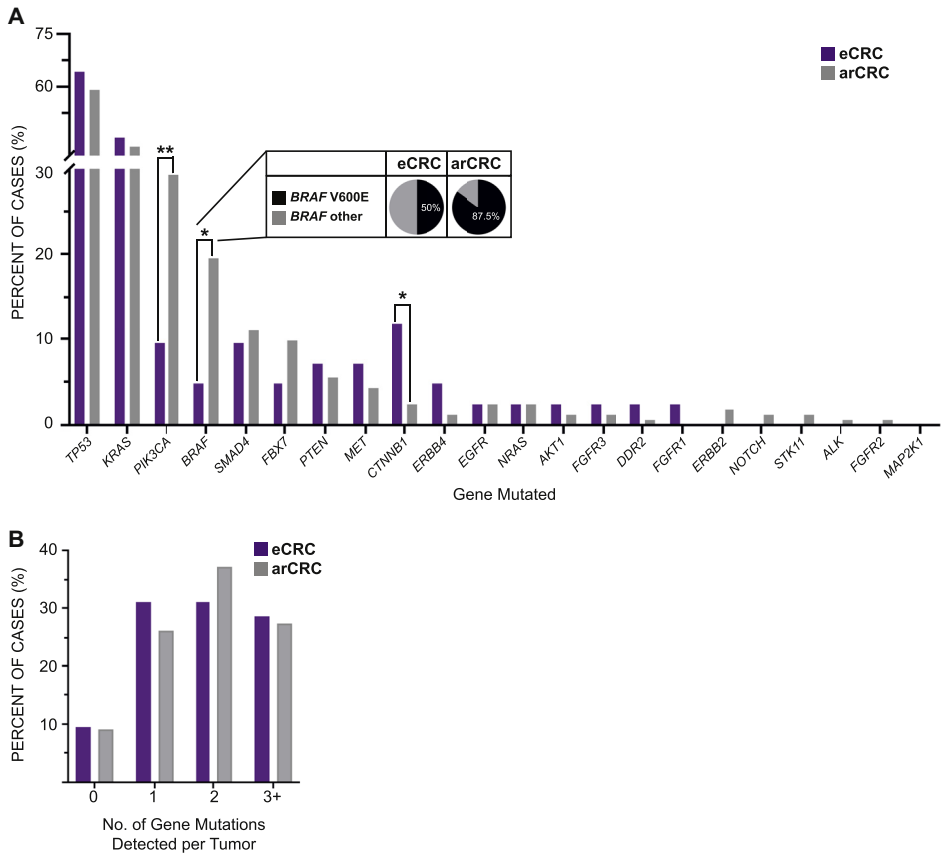

Fig. 1 Frequencies of genes mutated in eCRC and arCRC. **(A)** The most frequently mutated genes (overall frequencies in parenthesis) in eCRC and arCRC were *TP53* (60.2%) *KRAS* (43.2%), *PIK3CA* (25.3%), *BRAF* (16.5%) [inset shows frequency of *BRAF* V600E as a subset of total *BRAF* mutations (p = 0.012), *SMAD4* (10.7%), *FBX7* (8.7%), *PTEN* (5.8%), *MET* (4.9%), and *CTNNB1* (4.4%) (* = $P < 0.05$, ** = $P < 0.01$). **(B)** Number of gene mutations detected per tumor.

with arCRC (11.9% versus 2.4%, $P = 0.019$, Table 2). There was no significant difference in the rates of *TP53* mutation when comparing eCRC to arCRC.

Although the rates of MSS and MSI-H were similar between the 2 cohorts (Table 2), the genetic alteration pattern was distinct when comparing MSS or MSI-H tumors between the 2 cohorts. The frequency of *PIK3CA* mutations was significantly lower in MSS eCRC compared with MSS arCRC (5.7% versus 25.4%, Fig. 2A), with no difference in *PIK3CA* mutation frequency in MSI-H tumors between the 2 cohorts. While there was no difference in frequency of *KRAS* mutations at the cohort level (45.2% in eCRC compared to 42.6% in arCRC, Table 2), when stratifying by microsatellite instability status, we found that there was a significant increase in *KRAS* mutation frequency in eCRC MSI-H tumors (71.4%) versus arCRC MSI-H tumors (13.3%) ($P = 0.005$, Fig. 2B & Table 2). Of these, 100% of *KRAS* mutations in the eCRC MSI-H tumors involved codons 12 and 13. In comparison, 2 out of 4 *KRAS* mutations in arCRC MSI-H tumors were in codons 12 and 13, and 1 of 4 was an activating mutation at codon 61. MSI-H eCRC tumors were significantly more likely to harbor *CTNNB1* mutations (57.1% versus 3.3%, $P = 0.003$, Fig. 2C), whereas MSS tumors showed no difference in CTNNB1 mutation frequency between the 2 age cohorts.

The *BRAF* mutational landscape was significantly different between the eCRC and arCRC cohorts. MSS tumors in both cohorts harbored a low frequency of *BRAF* mutations (5.7% in eCRC and 7.5% in arCRC, Fig. 2D). While the frequency of *BRAF V600E* mutations in arCRC MSI-H tumors was high (73.3%), there were zero cases of eCRC MSI-H tumors that showed any *BRAF* mutations (Fig. 2D and Table 2). Loss of MLH1 in the absence of *BRAF* V600E mutation is, a combination that is suspicious for Lynch Syndrome, was found in 85.7% of eCRC MSI-H tumors (85.7%) compared with 6 of 30 arCRC MSI-H tumors in arCRC (20.0%) (Supplementary Table 1). Only one patient was known to have subsequently undergone germline genetic testing and was confirmed to have Lynch syndrome; this patient's tumor showed two mutations (one in *KRAS* and one in *TP53*). In the eCRC cohort, MLH1-deficient tumors showed high rates of mutation in *KRAS* (71.4% versus 10.0% in arCRC) and *CTNNB1* (57.1% versus 3.3% in arCRC) (Supplementary Table 1). MSH2 loss was a relatively rare event in our cohorts. One eCRC tumor with MSH2 loss harbored an *EGFR* mutation, whereas arCRC tumors with MSH2 loss showed single

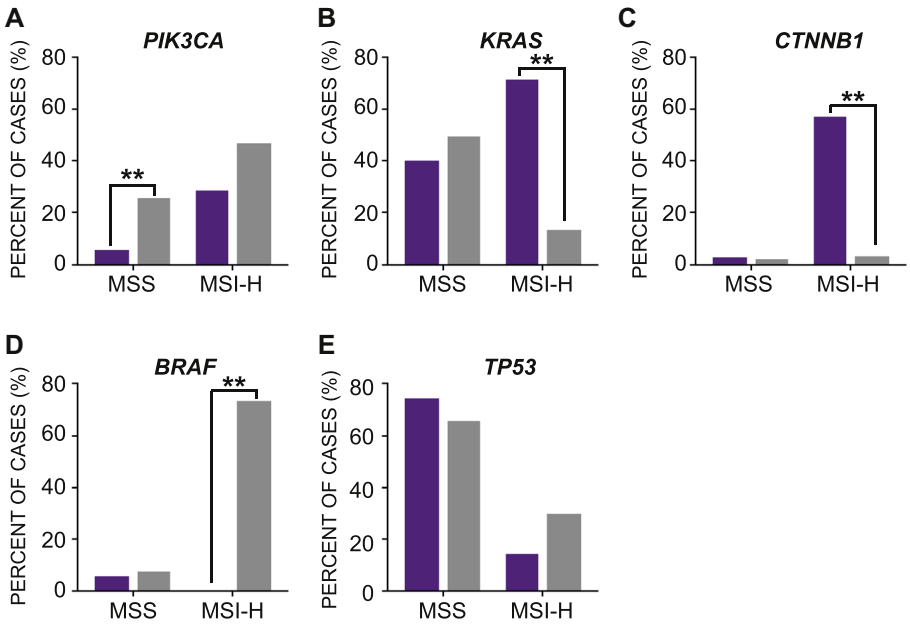

Fig. 2 Number of gene mutations per tumor in eCRC and arCRC. **(A)** *PIK3CA*, MSS; 5.7% eCRC versus 25.4% arCRC, $P = 0.010$, MSI-H; 28.6% eCRC versus 46.7% arCRC. **(B)** *KRAS*, MSS; 40.0% eCRC versus 49.3% arCRC, MSI-H; 71.4% eCRC versus 13.3% arCRC, $P = 0.005$. **(C)** *CTNNB1*, MSS; 2.9% eCRC versus 2.2% arCRC, MSI-H; 57.1% eCRC versus 3.3% arCRC, $P = 0.003$. **(D)** *BRAF*, MSS; 5.7% eCRC versus 7.5% arCRC, MSI-H; 0% eCRC versus 70.0% arCRC, $P = 0.001$. **(E)** *TP53*, MSS; 74.3% eCRC versus 65.7% arCRC, MSI-H; 14.3% eCRC versus 30.0% arCRC (** $= P < 0.01$). Abbreviation: MSS, microsatellite stable; MSI-H, microsatellite instability high.

mutations in *KRAS*, *PTEN*, *TP53*, and *NOTCH1* (Supplementary Table 1). No significant difference in the rates of *TP53* mutation was observed comparing eCRC and arCRC stratified by microsatellite instability status (Fig. 2E).

3.4. Clinical outcomes

Clinical follow-up data were available for 29 out of 42 cases in the eCRC cohort (69.0%) and for 130 out of 164 cases in the arCRC cohort (79.3%) and is shown in Supplementary Table 4. Neoadjuvant chemotherapy or radiation was reported in 21.4% eCRC patients, compared with 6.1% of arCRC patients. The rate of surgical resection was similar between the 2 cohorts (70.0% in eCRC and 76.2% in arCRC). Postsurgical adjuvant chemotherapy was given to 16.7% of the eCRC patients and 6.1% of the arCRC patients. One patient in the arCRC cohort received adjuvant radiation therapy. Two relapses were reported in the eCRC cohort, and 13 relapses were reported in the arCRC cohort during the surveillance time period, with a median time to relapse of 107 weeks and 90 weeks, respectively. One death was reported in the eCRC cohort after 29 weeks, and 7 deaths were reported in the arCRC cohort with a median survival time of 77 weeks. At the time of submission of this manuscript, the median survival time in the eCRC cohort was 220 weeks compared with 190 weeks in the arCRC cohort.

4. Discussion

Overall, these data show that early onset colorectal carcinoma is distinct in both clinicopathologic and molecular features from age-related colorectal carcinoma. Currently, there are no apparent agreed-upon definition of early age in colorectal carcinoma; the literature includes studies with a variety of definitions, some including patients in the 40−50 year old age range [21]. However, the trends in the United States and globally show an increased incidence of CRC between the ages of 40 and 50 years. Our approach was to restrict the definition of early-onset colorectal carcinoma and of age-related colorectal carcinoma, excluding patients between the ages of 40 and 70 years of age to diminish the co-mingling of eCRC and arCRC entities in our cohorts [22].

We found that eCRC is more likely to present with higher TNM stage and to be left sided or rectal in location, compared with arCRC which is more likely to present at lower TNM stage and more likely to be right-sided. While some studies have suggested that left-sided versus right-sided CRCs have different protein expression and genetic alteration profiles of *KRAS* and *BRAF* [23,24], a more recent study found no difference in *KRAS* or *BRAF* mutation and MMR protein expression between left-sided and right sided tumors [25]. These studies did not stratify cases by patient age of onset. Our results suggest that location and age of onset may be considered together as having unique pathologic and molecular features.

We determined MMR protein loss via an immunohistochemical approach, with greater than 1% of tumor cells showing bright intensity to be positive for protein expression. Recent reviews of the field of MMR IHC have shown that while cutoffs of 0% for loss of expression and >10% for retained expression correlate well with molecular tests confirming genetic mutation in the genes encoding MMR proteins, tumors with <10% of cells showing positivity are indeterminate for loss of expression/genetic mutation [13]. Further, the presence of intratumoral lymphocytes, which may be used as a positive internal control, may contribute ~1% of positive signal and is a potential pitfall in interpreting MMR IHC. Our institutional approach has been to report >1% of tumor cells with bright intensity as positive to avoid Type I error when reporting on loss of MMR protein expression.

In contrast to some studies showing higher rates of MSI-H tumors in early-onset colorectal cancer in contrast to age-related CRC [22,26], we found overall similar rates of MSS and MSI-H tumors in the eCRC and arCRC cohorts. This finding may reflect our more restrictive criteria for both early-onset and age-related CRC entities. Additionally, this may reflect the population of patients served at our institution, a tertiary care center in a large metropolitan setting in the Midwest United States. However, stratifying the data by age and microsatellite status reveals key differences between eCRC and arCRC. *PIK3CA* was less frequently altered in MSS eCRC, in contrast to some recent reports showing more *PIK3CA* mutations in early onset MSS tumors [22]. No difference in frequency of *PIK3CA* mutations was observed in MSI-H tumors between the 2 age cohorts. We found that while *KRAS* mutational frequency is similar in MSS tumors between the 2 age cohorts, MSI-H tumors in eCRC show a distinctly higher rate of *KRAS* mutation compared with MSI-H tumors in arCRC, confirming the findings of Serebriiskii et al. [27]. Furthermore, all of the *KRAS* mutations identified in the eCRC MSI-H tumor group involved codons 12 and 13 known to confer resistance to anti-EGFR therapy. This has important implications for treatment of eCRC, as the Ras GTPases Kras and Nras are important mediators of the EGFR pathway targeted by the successful antibody therapies cetuximab and panitumumab [28]. There was a very low frequency of *EGFR* mutation in both the eCRC and arCRC cohorts, regardless of MSI status.

In our cohort, *TP53* was the most frequently mutated gene in both the eCRC and arCRC. Although the frequency of *TP53* mutations in MSI-H tumors in the eCRC cohort tended to be lower, we did not identify a significant difference in frequency of *TP53* mutation in the early-onset versus age-related cohorts, regardless of MSI-status. This finding may reflect limitations in the size of our single-institution cohort to detect differences in the rate of *TP53* mutations by age.

Lynch syndrome (LS) is the most inherited cancer susceptibility syndrome, predisposing affected individuals to colorectal and endometrial carcinoma and less commonly the ovary, stomach, and other sites. Detection of pathogenic variants in DNA MMR including *MLH1, MSH2, MSH6, and PMS2* are the cornerstone of the diagnosis of LS. The paired protein complexes MSH2/MSH6 and MLH1/PMS2 are critical for both activity and downstream signaling by the MMR cellular pathway; MMR-deficient cells do not recruit ataxia-telangiectasia mutated (ATM) or ATM and Rad3-related proteins with subsequent impairment of p53-mediated cell cycle arrest and apoptosis [29]. Detection of MMR by immunohistochemistry has similar sensitivity to detection by microsatellite instability by molecular methods and is the preferred method for biopsies with low tumor cell proportion [30]. *MLH1*-deficient tumors that also lack *BRAF* V600E mutation are more suspicious for LS. In our cohort, we observed the combined loss of expression of hMLH1 in the absence of *BRAF* V600E, suspicious for LS, in 85.7% of MSI-H tumors in eCRC, whereas only 20.0% of MSI-H tumors in the arCRC cohort showed that combination. These findings support the concept that MSI-H status in eCRC is more highly correlated to LS compared with arCRC. Only one patient in the eCRC cohort, who had an MSI-H tumor, was known to have undergone germline genetic testing which confirmed the presence of LS. This patient's tumor showed 2 mutations, one in *KRAS* and one in *TP53*. To our knowledge, no other patients underwent germline genetic testing, and therefore, the rates of germline *APC*, *MUTYH/NTLH1, and TP53* mutations (associated with autosomal dominant FAP, autosomal recessive FAP, and Li-Fraumeni syndromes, respectively) in our cohorts is unknown. Furthermore, *APC* and *MUTYH* were not included in the 22-gene NGS panel in use at our institution over the time period of the study. The frequency of these mutations in eCRC, as well as the prevalence of familial cancer predisposition syndromes, remains a topic for further investigation.

Mutations in the *WNT/APC* pathway are known to have prognostic significance in colorectal cancer [26,31]. Recent literature shows upregulation of the *WNT/APC/CTNNB1* pathway, detected by increased nuclear beta-catenin expression, in early onset colorectal carcinoma [22,32]. However, *APC* mutations are decreased in the young [5]. Our study supports increased alteration rates of *CTNNB1* in eCRC, regardless of microsatellite instability status. A proposed molecular classification scheme for CRC [33] includes one subgroup of tumors that show chromosomal instability, without CpG island methylator phenotype and upregulation of the Wnt/beta-catenin pathway. Our data suggest that molecular genetic alteration of *CTNNB1* is a predominant feature of eCRC, especially MSI-H eCRC, being found in greater than 50% of such cases. While *FBX7*, a tumor suppressor protein that acts to antagonize Wnt signaling by targeting beta-catenin for degradation [34] is included in our 22 gene panel, we did not find any significant difference in *FBX7* mutational frequency between eCRC and arCRC regardless of MSI status. One

limitation of our study is that additional gene family members of the *WNT/APC* pathway are not included on our panel. Further work is needed to fully elucidate the contributions of this pathway to early-onset colorectal cancer.

Clinical outcomes data suggest that eCRC patients in our study experienced a longer median time to relapse and higher rates of overall survival compared with arCRC patients in our cohort. However, our study was not sufficiently powered to draw definitive conclusions about the clinical outcomes data between the cohorts. Additionally, the potential contribution of co-morbidities in an elderly population compared with a young population is not elucidated in the present study. Further clinical studies are needed to address outcomes and survival in early-onset versus late-onset colorectal carcinoma.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humpath.2020.08.002.

Acknowledgements

This study was supported in part by NIH R01 DK10776 (G-Y Yang), CA172431 (G-Y Yang), and CA164041 (G-Y Yang), as well as by the Harold E. Eisenberg Foundation Research Scholar Award (G-Y Yang).

References

[1] Society AC. Colorectal cancer facts & figures 2020-2022. Atlanta: American Cancer Society; 2020.
[2] Siegel RL, Torre LA, Soerjomataram I, et al. Global patterns and trends in colorectal cancer incidence in young adults. Gut 2019;68:2179−85.
[3] Siegel RL, Fedewa SA, Anderson WF, et al. Colorectal cancer incidence patterns in the United States, 1974-2013. J Natl Canc Inst 2017:109.
[4] Meyer JE, Cohen SJ, Ruth KJ, Sigurdson ER, Hall MJ. Young age increases risk of lymph node positivity in early-stage rectal cancer. J Natl Canc Inst 2016;108.
[5] Lieu CH, Renfro LA, Gramont Ad, et al. Association of age with survival in patients with metastatic colorectal cancer: analysis from the ARCAD clinical trials program. J Clin Oncol 2014;32:2975−82.
[6] Goel A, Nagasaka T, Spiegel J, Meyer R, Lichliter WE, Boland CR. Low frequency of Lynch syndrome among young patients with non-familial colorectal cancer. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2010;8:966−71.
[7] Alsop K, Mead L, Smith LD, et al. Low somatic K-ras mutation frequency in colorectal cancer diagnosed under the age of 45 years. EJC (Eur J Canc) 2006;42:1357−61.
[8] Pearlman R, Frankel WL, Swanson B, et al. Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. JAMA Oncol 2017;3:464−71.
[9] Lutgens MW, Vleggaar FP, Schipper ME, et al. High frequency of early colorectal cancer in inflammatory bowel disease. Gut 2008;57:1246−51.
[10] Willauer AN, Liu Y, Pereira AAL, et al. Clinical and molecular characterization of early-onset colorectal cancer. Cancer 2019;125:2002−10.
[11] Chang DT, Pai RK, Rybicki LA, et al. Clinicopathologic and molecular features of sporadic early-onset colorectal adenocarcinoma: an adenocarcinoma with frequent signet ring cell differentiation, rectal and sigmoid involvement, and adverse morphologic features. Mod Pathol Off J Unit States Canad Acad Pathol 2012;25:1128−39. Inc.
[12] Shia J. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary non-polyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. J Mol Diagn 2008;10:293−300.
[13] Sarode VR, Robinson L. Screening for Lynch syndrome by immunohistochemistry of mismatch repair proteins: significance of indeterminate result and correlation with mutational studies. Arch Pathol Lab Med 2019;143:1225−33.
[14] Deeken-Draisey A, Yang G-Y, Gao J, Alexie BA. Anaplastic thyroid carcinoma: an epidemiologic, histologic, immunohistochemical and molecular single institution study. Hum Pathol 2018;82:140−8.
[15] Consortium O. Ion AmpliSeq colon and lung cancer research panel v2 and Ion AmpliSeq RNA fusion lung cancer research panel. ThermoFish Sci 2016:4.
[16] D'Haene N, Le Mercier M, De Neve N, et al. Clinical validation of targeted next generation sequencing for colon and lung cancers. PloS One 2015;10:e0138245.
[17] Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol 2011;29:24−6.
[18] Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings Bioinf 2013;14:178−92.
[19] Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 2017;45:D777−83.
[20] Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inf 2019;95:103208.
[21] Fairley TL, Cardinez CJ, Martin J, et al. Colorectal cancer in U.S. adults younger than 50 years of age, 1998−2001. Cancer 2006;107:1153−61.
[22] Kirzin S, Marisa L, Guimbaud R, et al. Sporadic early-onset colorectal cancer is a specific sub-type of cancer: a morphological, molecular and genetics study. PloS One 2014;9:e103159.
[23] Tong JHM, Lung RWM, Sin FMC, et al. Characterization of rare transforming KRAS mutations in sporadic colorectal cancer. Canc Biol Ther 2014;15:768−76.
[24] Brulé SY, Jonker DJ, Karapetis CS, et al. Location of colon cancer (right-sided versus left-sided) as a prognostic factor and a predictor of benefit from cetuximab in NCIC CO.17. Eur J Canc 2015;51:1405−14.
[25] Gao XH, Yu GY, Gong HF, et al. Differences of protein expression profiles, KRAS and BRAF mutation, and prognosis in right-sided colon, left-sided colon and rectal cancer. Sci Rep-uk 2017;7:7882.
[26] Lieu CH, Golemis EA, Serebriiskii IG, et al. Comprehensive genomic landscapes in early and later onset colorectal cancer. Clin Canc Res Off J Am Assoc Canc Res 2019;25:5852−8.
[27] Serebriiskii IG, Connelly C, Frampton G, et al. Comprehensive characterization of RAS mutations in colon and rectal cancers in old and young patients. Nat Commun 2019;10:3722.
[28] Amado RG, Wolf M, Peeters M, et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. J Clin Oncol 2008;26:1626−34.
[29] Brown KD, Rathi A, Kamath R, et al. The mismatch repair system is required for S-phase checkpoint activation. Nat Genet 2003;33:80−4.
[30] Cerretelli G, Ager A, Arends MJ, Frayling IM. Molecular pathology of Lynch syndrome. J Pathol 2020;250:518−31.

[31] Schell MJ, Yang M, Teer JK, et al. A multigene mutation classification of 468 colorectal cancers reveals a prognostic role for APC. Nat Commun 2016;7:11743.

[32] Voer R, Hahn M-M, Weren RDA, et al. Identification of novel candidate genes for early-onset colorectal cancer susceptibility. PLoS Genet 2016;12:e1005880.

[33] Marisa L, Reyniès Ad, Duval A, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. PLoS Med 2013;10:e1001453.

[34] Jiang J, Sun C, Tian S, Yu C, Chen M, Zhang H. Tumor suppressor Fbxw7 antagonizes WNT signaling by targeting β-catenin for degradation in pancreatic cancer. Tumor Biol 2016;37:13893–902.