Heliyon 10 (2024) e29299

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Correlation between NGS panel-based mutation results and clinical information in colorectal cancer patients

Bo Cheng [a,1], Lin Xu [a,1], Yunzhi Zhang [b,c,1], Huimin Yang [b], Shan Liu [a], Shanshan Ding [a], Huan Zhao [a], Yi Sui [b], Chan Wang [b], Lanju Quan [a], Jinhong Liu [a], Ye Liu [a], Hongming Wang [b], Zhaoqing Zheng [b], Xizhao Wu [a], Jing Guo [a], Zhaohong Wen [b], Ruya Zhang [a], Fei Wang [a], Hongmei Liu [b,**], Suozhu Sun [a,*]

[a] *Department of Pathology, Chinese People's Liberation Army Rocket Force Characteristic Medical Center, Beijing 100037, China*
[b] *Singlera Genomics (Shanghai) Ltd., Shanghai 201318, China*
[c] *School of Life Sciences, Fudan University, Shanghai 200438, China*

ARTICLE INFO

Keywords:
Colorectal cancer
Next-generation sequencing
gene mutation
microsatellite instability
Molecular targeted therapy

ABSTRACT

Early mutation identification guides patients with colorectal cancer (CRC) toward targeted therapies. In the present study, 414 patients with CRC were enrolled, and amplicon-based targeted next-generation sequencing (NGS) was then performed to detect genomic alterations within the 73 cancer-related genes in the OncoAim panel. The overall mutation rate was 91.5 % (379/414). Gene mutations were detected in 38/73 genes tested. The most frequently mutated genes were *TP53* (60.9 %), *KRAS* (46.6 %), *APC* (30.4 %), *PIK3CA* (15.9 %), *FBXW7* (8.2 %), *SMAD4* (6.8 %), *BRAF* (6.5 %), and *NRAS* (3.9 %). Compared with the wild type, *TP53* mutations were associated with low microsatellite instability/microsatellite stability (MSI-L/MSS) ($P = 0.007$), tumor location ($P = 0.043$), and histological grade ($P = 0.0009$); *KRAS* mutations were associated with female gender ($P = 0.026$), distant metastasis ($P = 0.023$), TNM stage ($P = 0.013$), and histological grade ($P = 0.004$); *APC* mutations were associated with patients <64 years of age at diagnosis ($P = 0.04$); *PIK3CA* mutations were associated with tumor location ($P = 4.97e\text{-}06$) and female gender ($P = 0.018$); *SMAD4* mutations were associated with tumor location ($P = 0.033$); *BRAF* mutations were associated with high MSI (MSI-H; $P = 6.968e\text{-}07$), tumor location ($P = 1.58e\text{-}06$), and histological grade ($P = 0.04$). Mutations in 164 individuals were found to be pathogenic or likely pathogenic. A total of 26 patients harbored MSI-H tumors and they all had at least one detected gene mutation. Mutated genes were enriched in signaling pathways associated with CRC. The present findings have important implications for improving the personalized treatment of patients with CRC in China.

1. Introduction

Colorectal cancer (CRC), one of the most common and deadly malignancies worldwide, is caused by genetic events and epigenetic

* Corresponding author.
** Corresponding author.
E-mail addresses: hmliu1989@126.com (H. Liu), mingzhu0368@sina.com (S. Sun).
[1] These authors contributed to the work equally.

https://doi.org/10.1016/j.heliyon.2024.e29299
Received 15 December 2023; Received in revised form 4 April 2024; Accepted 4 April 2024
Available online 5 April 2024

alterations. The advent of the era of precision medicine has brought good news to patients with CRC [1]. Genetic and gut microbiota have proven to be possible causes of CRC [2]. The National Comprehensive Cancer Network (NCCN) guidelines proposed individual genetic tests that could be used to assess hereditary CRC susceptibility. *HER2*, *KRAS*, *NRAS*, and *BRAF* testing has shown good performance in targeting treatment [3], while *APC*, *TP53*, *PTEN*, and *PIK3CA* genes have been found to be associated with the diagnosis, treatment, and prognosis of CRC [3–5]. Microsatellite instability (MSI) and deficient mismatch repair (dMMR) are emerging issues in the oncology and molecular pathology fields [6]. MSI/dMMR is important for prognosis prediction, anticancer response, and immunotherapy in metastatic and non-metastatic tumors [7,8]. Furthermore, high MSI (MSI-H) associated with downstream frameshift mutations is a biomarker that guides immunotherapy for CRC [9–11]. Understanding the relationship between genetic mutation signatures, pathological classification, and tumor stage can lead to better prognosis and treatment for CRC [3]. The Cancer Genome Atlas project provides a comprehensive view of tumor genomic changes for CRC and confirms the role of vital drive genes such as *SMAD4*, *KRAS*, *PIK3CA*, *TP53*, and *APC* in tumorigenesis [9,12]. *APC* and *TP53,* as diagnosis signatures, are the most common mutation combinations. Activating or losing them gradually promotes the transformation of adenomas into cancer. As a crucial signaling pathway in CRC, *KRAS* and *BRAF* mutations are associated with distant metastases and poor prognosis [13–15]. Mutations in the *TP53* and *SMAD4* genes promote the transition of adenoma to invasive carcinoma [16,17]. Studies have shown that *APC*, *KRAS*, and *BRAF* mutations are considered early events in CRC development [14,18]. Nevertheless, the role of driver genes in CRC metastasis remains to be further elucidated.

Next-generation sequencing (NGS) has several potential advantages, such as increasing the overall accuracy of variant calling, while probe dropout and inadequate coverage may be problems with exome sequencing methods [19]. Furthermore, the identification of fusions [20] and other structural variants [21] (by detection of split reads) now yields a higher accuracy than alternative techniques [22].

The aim of the present study was to determine the molecular spectrum of 73 cancer-related gene mutations using the OncoAim panel, which adopted an amplification-based enrichment method that has been proven more successful than probe-based capture approach in FFPE tissue, requiring lower DNA input). The panel was accompanied with bioinformatics pipeline optimized for somatic variant detection and covered more than 6000 hotspots in 59 cancer genes and 14 chemotherapy-related genes included in Pharmacogenetics and Pharmacogenomics Knowledge Base (PharmGKB) [23,24]. Clinical characteristics, such as age, gender, anatomic site, TNM staging of tumors, distant metastasis or lymph node metastasis, smoking history, alcohol history, and MSI status, were collected from a group of patients with CRC at the Chinese PLA Rocket Force General Hospital. Genetic alterations were studied alongside clinical and pathological factors to identify diagnostically important mutations that could potentially guide targeted therapy.

2. Methods

2.1. Patient recruitment

A total of 414 patients were included according to standard procedures. Samples were collected in Rocket Army Specialized Medical Center, Beijing, China. All the patients enrolled had been histologically confirmed as CRC, according to World Health Organization criteria and reviewed by professional pathologists, based on hematoxylin and eosin staining. This study was approved by the Ethics Committee of the Chinese People's Liberation Army Rocket Force Characteristic Medical Center (Approval no.KY2023038) and conducted according to the principles of the Declaration of Helsinki [25].

2.2. DNA extraction from formalin-fixed, paraffin-embedded (FFPE) samples

All tissue sections were reviewed by pathologists to ensure that tumor cell content is >10 %. DNA was isolated from FFPE tissues using the QIAamp DNA FFPE Tissue Kit (Cat. No. 56404; Qiagen), following the manufacturer's instructions. Briefly, paraffin was first dissolved in xylene and then removed. The samples were then lysed under denaturing conditions and briefly digested with proteinase K. The samples were incubated at 90 °C to reverse the formalin cross-linking process. The DNA was then able to bind to the membrane, after which any contaminants were washed away. The DNA was finally eluted in Buffer ATE, and it can be used immediately for amplification reactions or stored at −20 °C. The concentrations were detected using the Qubit Fluorometer (Cat. No. Q33238; Thermo Fisher Scientific). Only the DNA samples with most fragments over 500 bp and >50 ng could be included in subsequent experiments.

2.3. Next-generation sequencing and data analysis

Amplicon-based targeted NGS was performed on 414 tumor samples using the OncoAim® kit (Cat. No. A01D; Singlera Genomics, Inc.), which consists of all the exons of 73 CRC-related genes (Supplementary table 1). Following DNA concentration detection, the input of FFPE DNA for library preparation was 50 ng for each sample. Accordingly, DNA was sheared to ~250 bp prior to library construction. Following end repair, A-tailing and adapter ligation, target capture using the kit's included probes was carried out, following the manufacturer's instructions. Using 150 bp paired-end runs on the NextSeq 500 (Cat. No. SY-415-1001; Illumina, Inc.), the library product was sequenced. The overall median coverage of depth was >1,000X and the uniformity was >90 %.

The OncoAim® kit, which can detect single nucleotide variants, and short insertions and deletions (InDels) at the same instant, was used to process the sequencing data. Briefly, the sequencing reads in the FASTQ format were qualified and filtered using Trim Galore (version 0.4.0), and then aligned to the human genome reference sequence GRCh37/hg19 using Burrows-Wheeler Aligner (version

0.7.12). The ENCODE Blacklist was used to filter the aberrant high-signal regions in the genome. Following mapping, freebayes (version 1.0.2) was used to call single nucleotide variations, insertions, and deletions. Human Genome Variation Society notation for single nucleotide variations, insertions, and deletions was performed using Ensembl Variant Effect Predictor (release-84) and a self-developed flow based on the ClinVar, dbSNP, and COSMIC databases. The gene variation results were classified as pathogenic, likely pathogenic, benign, likely benign, uncertain or inconclusive, according to information from the ClinVar database. The minimum confidence threshold for variants was 2 %. In particular, for variations with a frequency of 2–5 %, manual reviews were performed by Integrative Genomics Viewer (IGV) and the confident variations were ultimately reported.

2.4. MSI testing

Both the National Cancer Institute (NCI) 2B3D (BAT-25, BAT-26, D2S123, D17S250, and D5S346) and the single nucleotide site (BAT-25, BAT-26, NR-21, NR-24, NR-27, and MONO-27) panels based on fluorescence quantitative PCR and capillary electrophoresis were used for primary and metastatic tumor tissue MSI detection. DNA products were used as templates for PCR amplification. The 2B3D panel was built at our laboratory (Department of Pathology, Chinese People's Liberation Army Rocket Force Characteristic Medical Center, China) with reference to the nucleotide sites and primer sequences recommended by NCI [26], and the PCR products were analyzed by capillary electrophoresis using the Applied Biosystems 3500 Dx Genetic Analyzer (Cat. No. 4461450; Thermo Fisher Scientific). The single nucleotide site panel was detected by the Microsatellite Unstable Gene Detection Kit (Cat. No. 20213400936; Beijing Microread Genetics Co., Ltd.), following the manufacturer's instructions, and the PCR products were detected by capillary electrophoresis using the GenReader 7010 Genetic Analyzer (Cat. No. 20212220099; Beijing Microread Genetics Co., Ltd.). Based on the detection outcome, samples with two or more unstable sites were identified as MSI-H, those with one unstable site were identified as MSI-L, and those with no unstable sites were identified as microsatellite stability (MSS). If the results of the two panels were consistent, they were recorded as experimental results, otherwise they were recorded as inconsistent.

Table 1
Clinical characteristics of enrolled patients with colorectal cancer.

Characteristic	No. of patients (%)
Age: Median [range], y	64 [56–71]
<64	191 (46.1)
≥64	223 (53.9)
Gender	
Men	275 (66.4)
Women	139 (33.6)
Anatomic site	
Rectum	237 (57.2)
Left colon	72 (17.4)
Right colon	67 (16.2)
Multiple or unclear	38 (9.2)
Initial stage at diagnosis	
0	10 (2.4)
I	81 (19.6)
II	118 (28.5)
III	110 (26.6)
IV	36 (8.7)
Unknown	59 (14.2)
Distant metastasis	
No	345 (83.3)
Yes	58 (14.0)
Histological grade	
Poor differentiation	12 (2.9)
Medium differentiation	368 (88.9)
Well differentiation	21 (5.1)
Lymph node metastasis	
No	254 (61.4)
Yes	160 (38.6)
Smoking history	
None	315 (76.1)
Current	96 (23.2)
Former	3 (0.7)
Alcohol history	
No	299 (72.2)
Yes	107 (25.8)
Microsatellite instability	
MSI-H	26 (6.28)
MSI-L/MSS	367 (88.65)

2.5. Statistical analysis

R statistical programming language (R Core Team [2022], version 4.0.4) was used for the statistical study. To examine the relationship between the mutational status and the demographic and clinical factors, the χ^2-square or Fisher's exact test was performed. $P < 0.05$ was considered to indicate a statistically significant difference.

3. Results

3.1. Patient baseline information

A total of 414 patients with CRC [275 (66.4 %) men; 139 (33.6 %) women] were enrolled in the present study, of which 379 patients had at least one gene mutation. The median age at initial diagnosis was 64 years (range, 56–71 years). There were 67 (16.2 %) cases of right-sided colon cancer, 72 (17.4 %) cases of left-sided colon cancer, and 237 (57.2 %) cases of rectal cancer. There were also 38 (9.2 %) patients with CRC with multiple or unclear tumor locations. There were 10 (2.4 %) stage 0, 81 (19.6 %) stage I, 118 (28.5 %)

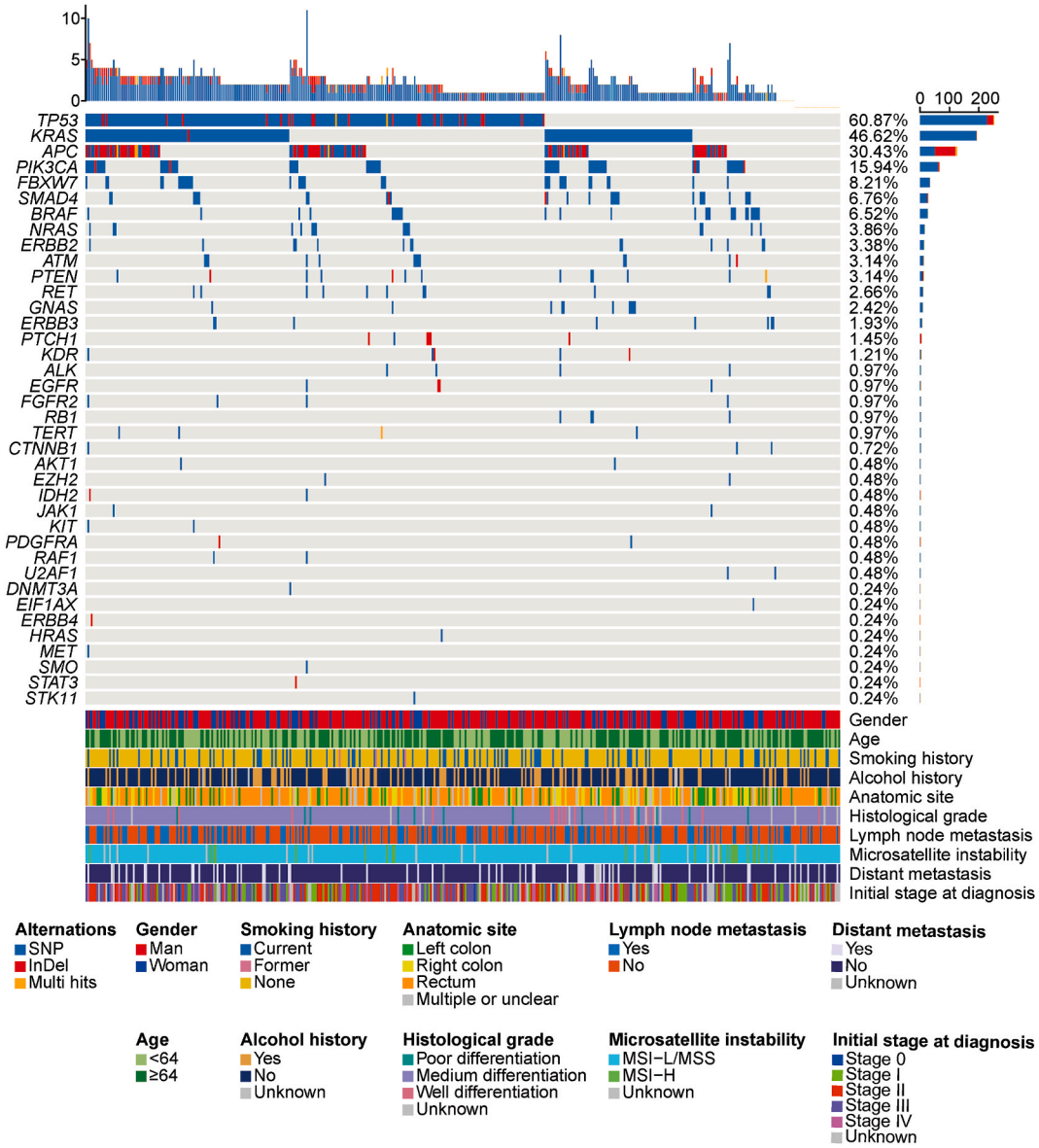

Fig. 1. Mutational landscape and clinical characteristics of 414 patients with CRC. A total of 38 genes with mutations are shown. The x axis represents 414 patients, and the y axis represents the mutant genes and the corresponding mutation frequency. Clinical characteristics are annotated in the bottom.

stage II, 110 (26.6 %) stage III, and 36 (8.7 %) stage IV tumors. MSI analysis revealed that among the 379 patients with mutations, 26 were MSI-H and 333 were MSI-L/MSS. Additional clinical characteristics collected from the patients, such as histological grade, lymph node metastasis, distant metastasis, alcohol history, and smoking history, are shown in Table 1.

3.2. Genetic landscape of patients

The overall mutation rate was 91.5 % (379/414). Gene mutations were detected in 38 of the 73 genes tested. *TP53*, *KRAS*, *APC*, *PIK3CA*, *FBXW7*, *SMAD4*, *BRAF*, and *NRAS* mutations were found in 252 (60.9 %), 193 (46.6 %), 126 (30.4 %), 66 (15.9 %), 34 (8.2 %), 28 (6.8 %), 27 (6.5 %), and 16 (3.9 %) cases, highlighting the relatively high mutation rate in the study (Fig. 1). The correlation between total mutations and baseline data is shown in Table 2. The total mutation rate was independent of all the factors, including age, gender, MSI, distant metastasis, TNM stage, lymph node metastasis, histological grade, smoking history, tumor location or alcohol history.

In all mutation cases, mutations in 164 individuals were found to be pathogenic or likely pathogenic. In detail, we identified 1 patient with 6 pathogenic mutations, 3 patients with 5 pathogenic mutations, 5 patients with 4 pathogenic mutations, 15 patients with 3 pathogenic mutations, 75 patients with 2 pathogenic mutations, and 65 patients with 1 pathogenic mutation (Supplementary fig. 1). These pathogenic mutations suggested that these patients may have a poor prognosis [27].

3.3. The association between genotype and clinical characteristics

Next, patients were scrutinized for potentially actionable genetic alterations that may inform the selection of or predict response to targeted therapies. The clinical characteristics of 26 MSI-H patients are shown in the mutational landscape in Fig. 2. The result showed that 26 MSI-H patients had pathogenic or likely pathogenic mutations in the *TP53* (n = 9), *KRAS* (n = 7), *PIK3CA* (n = 8), *SMAD4* (n = 2), *BRAF* (n = 10), *NRAS* (n = 1), *FBXW7* (n = 1), and *APC* (n = 6) genes, while 367 MSI-L/MSS patients harbored alterations in the *TP53* (n = 232), *KRAS* (n = 174), *APC* (n = 109), *PIK3CA* (n = 56), *FBXW7* (n = 32), *SMAD4* (n = 23), *BRAF* (n = 16), and *NRAS* (n = 14) genes.

Compared with the wild type, TP53 mutations were associated with tumor location ($P = 0.043$), histological grade ($P = 0.0009$), and MSI-L/MSS (63.2 vs. 34.6 %; $P = 0.007$) (Table 3), with no significant correlation observed with other baseline data such as age, gender, distant metastasis, TNM stage, lymph node metastasis, smoking, and alcohol history. Compared with the wild type, *KRAS* mutations were associated with female gender (54.7 vs. 42.5 %; $P = 0.026$), distant metastasis (60.3 vs. 44.1 %; $P = 0.023$), TNM stage ($P = 0.013$), and histological grade ($P = 0.004$), and trended toward an association with MSI-L/MSS (47.4 vs 26.9 %, $P = 0.07$), tumor location ($P = 0.073$), <64 years of age at diagnosis (51.3 vs 42.6 %; $P = 0.09$), and smoking history ($P = 0.09$). There was no significant correlation between *KRAS* mutations and other baseline data. Compared with the wild type, *APC* mutations were associated with <64 years of age at diagnosis (35.6 vs 26 %; $P = 0.04$), with no significant correlation observed with other baseline data. Compared with the wild type, *PIK3CA* mutations were associated with location ($P = 4.97e\text{-}06$) and female gender (22.3 vs. 12.7 %; $P = 0.018$), trended toward an association with MSI-H (30.8 vs. 15.3 %; $P = 0.05$), and had no significant correlation with other baseline data. Compared with the wild type, *SMAD4* mutations were associated with location ($P = 0.033$), trended toward an association with lymph node (10.0 vs. 4.7 %; $P = 0.06$) and distant (12.1 vs. 5.8 %; $P = 0.09$) metastasis, and had no significant correlation with other baseline data. Compared with the wild type, *BRAF* mutations were associated with MSI-H (38.5 vs. 4.4 %; $P = 6.968e\text{-}07$), tumor location ($P = 1.58e$-

Table 2
Correlation analysis between total mutations and clinical characteristics.

Characteristic	Mutation cases	*P*-value
Age <64 y vs. ≥64 y	180 vs. 199	0.10
Gender Men vs. women	248 vs. 131	0.22
Anatomic site Left colon vs. right colon vs. rectum	65 vs. 66 vs. 211	0.056
Microsatellite instability MSI-H vs. MSI-L/MSS	26 vs. 333	0.15
Smoking history Current vs. none vs. former	88 vs. 288 vs. 3	1
Alcohol history Yes vs. No	101 vs. 270	0.27
Distant metastasis Yes vs. No	53 vs. 317	0.80
Initial stage at diagnosis 0 vs. I vs. II vs. III vs. IV	10 vs. 72 vs. 110 vs. 103 vs. 35	0.56
Histological grade Poor vs. medium vs. well differentiation	12 vs. 337 vs. 20	0.88
Lymph node metastasis Yes vs. No	150 vs. 229	0.27

MSI-H, high microsatellite instability; MSI-L, low microsatellite instability; MSS, microsatellite stability.

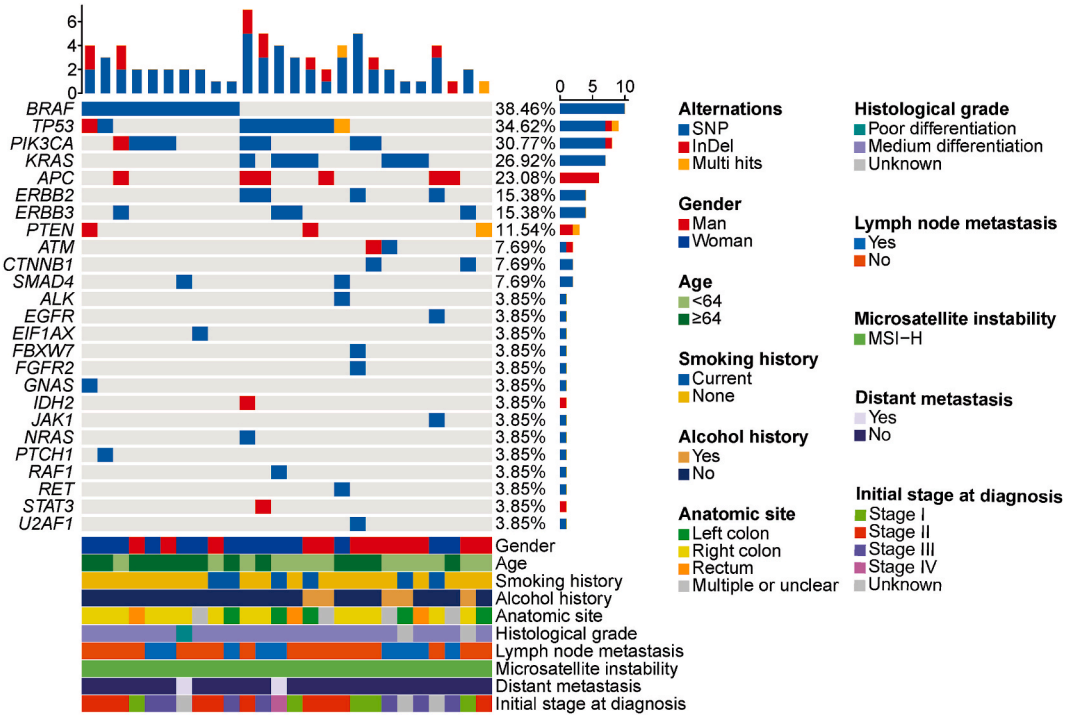

Fig. 2. Mutational landscape and clinical characteristics of 26 high MSI-H patients. A total of 25 genes with mutations are shown. The x axis represents 26 MSI-H patients, and the y axis represents the mutant genes and the corresponding mutation frequency. Clinical characteristics are annotated on the right.

06), and histological grade ($P = 0.04$), trended toward an association with patients without history of drinking (8 vs. 2.8 %; $P = 0.07$) and female gender (10.1 vs. 4.7 %; $P = 0.06$), and had no significant correlation with other baseline data. *NRAS* and *FBXW7* mutations were not associated with any baseline data. Finally, the above results indicated that the patients' drinking or smoking history and presence of lymph node metastasis were not significantly associated with mutations in the present study. Genotypes and clinical correlations are summarized in Table 3.

3.4. Pathway enrichment analysis of mutant genes

The major functions of the detected mutated genes were further analyzed using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. As shown in Fig. 3A and B, mutant genes were enriched in EGFR tyrosine kinase inhibitor resistance or CRC associated genes. In addition, the mutant genes were enriched in several carcinogenic signaling pathways, such as PI3K-Akt, MAPK, ErbB, Ras, VEGF, mTOR, JAK-STAT, and HIF-1 signaling pathways. The mutant oncogenes were also enriched in genes regulating PD-L1 expression and PD-1 checkpoint pathway in cancer, as well as several immune-related pathways, such as Fc epsilon RI, B cell receptor, and T cell receptor signaling pathways.

4. Discussion

In the present study, 414 patients with CRC were enrolled and NGS was performed to detect genomic alterations within the OncoAim panel of 73 cancer-related genes. The results showed that 91.5 % of tumors (379/414) had at least one mutation. Eight genes (*TP53, KRAS, APC, PIK3CA, FBXW7, SMAD4, BRAF*, and *NRAS*) were identified as highly mutated, and their correlation with clinical features was further explored. In addition, the MSI testing revealed that 26 patients harbored MSI-H tumors and 367 patients harbored MSI-L/MSS tumors. Of note, all MSI-H patients had at least one detected gene mutation. For the 38 genes with detected mutations, KEGG pathway enrichment analysis was further performed, and the result showed that genes regulating PD-L1 expression and PD-1 checkpoint pathway were enriched. The identification of diagnostically relevant genetic alterations will provide guidance for the development and application of personalized therapy for Chinese patients with CRC, thus improving future healthcare efficiency and resource utilization.

The gene mutations detected in the present study and their correlation with the clinical characteristics of patients with CRC were comprehensively discussed. *TP53*, *KRAS*, and *APC* mutations were the more frequent alterations observed among the eight genes in our cohort, which was consistent with the findings of previous studies [28–30]. In a sporadic study among young CRC patients, Abedalrhman Alkhateeb et al. found *KRAS* (G12V, G12D, A146T) and *APC* genes were pathogenic mutations. The present results indicated

Table 3
Correlation analysis between genotype mutations and clinical characteristics.

Characteristic	Cases of mutant patients (no.)/proportion (%)							
	TP53	KRAS	APC	PIK3CA	FBXW7	SMAD4	BRAF	NRAS
Age:	132 vs.	95 vs.	58 vs.	38 vs.	13 vs.	13 vs.	16 vs.	9 vs.
≥64 y vs.	120	98	68	28	21	15	11	7
<64 y	59.2 vs. 62.8	42.6 vs. 51.3	26.0 vs. 35.6	17.0 vs. 14.7	5.8 vs. 11.0	5.8 vs. 7.9	7.2 vs. 5.8	4.0 vs. 3.7
P-value	0.51	0.09	0.04[a]	0.60	0.08	0.53	0.70	1
Gender:	167 vs.	117 vs.	86 vs.	35 vs.	24 vs.	16 vs.	13 vs.	8 vs.
Men vs. women	85	76	40	31	10	12	14	8
	60.7 vs.	42.5 vs.	31.3 vs.	12.7 vs.	8.7 vs.	5.8 vs.	4.7 vs.	2.9 vs.
	61.2	54.7	28.8	22.3	7.2	8.6	10.1	5.8
P-value	1	0.026[a]	0.68	0.018[a]	0.73	0.38	0.06	0.25
Anatomic site:	155 vs.	112 vs.	65 vs.	23 vs.	19 vs.	12 vs.	8 vs.	11 vs.
Rectum vs. left	41 vs. 33	26 vs. 37	23 vs. 27	14 vs. 23	8 vs. 6	10 vs. 4	3 vs. 14	2 vs. 2
colon vs. right colon	65.4 vs. 56.9	47.3 vs. 36.1	27.4 vs. 31.9	9.7 vs. 19.4 vs.	8.0 vs.	5.1 vs. 13.9	3.4 vs. 4.2 vs.	4.6 vs. 2.8
	vs. 49.3	vs. 55.2	vs. 40.3	34.3	11.1 vs. 9.0	vs. 6.0	20.9	vs. 3.0
P-value	0.043[a]	0.073	0.125	<0.001[a]	0.718	0.033[a]	<0.001[a]	0.803
Microsatellite	9 vs.	7 vs.	6 vs.	8 vs.	1 vs.	2 vs.	10 vs.	1 vs.
instability:	232	174	109	56	32	23	16	14
MSI-H vs.	34.6 vs.	26.9 vs.	23.1 vs.	30.8 vs.	3.8 vs.	7.7 vs.	38.5 vs.	3.8 vs.
MSI-L/MSS	63.2	47.4	29.7	15.3	8.7	6.3	4.4	3.8
P-value	0.007[a]	0.07	0.62	0.05	0.71	0.68	<0.001[a]	1
Smoking history:	62 vs.	39 vs.	22 vs.	10 vs.	6 vs.	7 vs.	6 vs.	3 vs.
Current vs. none	187 vs. 3	154 vs. 0	103 vs. 1	55 vs. 1	28 vs. 0	21 vs. 0	21 vs. 0	12 vs. 1
vs. former	64.6 vs.	40.6 vs.	22.9 vs.	10.4 vs.	6.3 vs.	7.3 vs.	6.3 vs.	3.1 vs.
	59.4 vs.	48.9 vs.	32.7 vs.	17.5 vs.	8.9 vs.	6.7 vs.	6.7 vs.	3.8 vs.
	100.0	0.0	33.3	33.3	0.0	0.0	0.0	33.3
P-value	0.28	0.09	0.12	0.14	0.63	0.85	1	0.14
Alcohol history:	70 vs.	53 vs.	30 vs.	12 vs.	9 vs.	7 vs.	3 vs.	2 vs.
Yes vs. no	177	135	95	53	25	20	24	14
	65.4 vs.	49.5 vs.	28.0 vs.	11.2 vs.	8.4 vs.	6.5 vs.	2.8 vs.	1.9 vs.
	59.2	45.2	31.8	17.7	8.4	6.7	8.0	4.7
P-value	0.29	0.50	0.54	0.13	1	1	0.07	0.26
Distant metastasis:	34 vs.	35 vs.	17 vs.	7 vs.	3 vs.	7 vs.	3 vs.	1 vs.
Yes vs. no	215	152	107	57	31	20	24	14
	58.6 vs.	60.3 vs.	29.3 vs.	12.1 vs.	5.2 vs.	12.1 vs.	5.2 vs.	1.7 vs.
	62.3	44.1	31.0	16.5	9.0	5.8	7.0	4.1
P-value	0.66	0.023[a]	0.88	0.44	0.45	0.09	0.78	0.71
Initial stage at diagnosis:	6 vs. 43 vs. 77	8 vs. 34 vs.	2 vs. 23 vs.	2 vs. 7 vs. 25	2 vs. 4 vs.	0 vs. 2 vs. 6	0 vs. 2 vs. 7	0 vs. 4 vs. 6
0 vs. I vs. II	vs. 73	48	37	vs. 18	9	vs. 11	vs. 11	vs. 2
vs. III	vs. 25	vs. 53	vs. 34	vs. 6	vs. 12	vs. 4	vs. 2	vs. 1
vs. IV		vs. 24	vs. 10		vs. 0			
	60.0 vs. 53.1	80.0 vs. 42.0	20.0 vs. 28.4	20.0 vs. 8.6 vs.	20.0 vs.	0.0 vs. 2.5 vs.	0.0 vs. 2.5 vs.	0.0 vs. 4.9
	vs. 65.3 vs.	vs. 40.7 vs.	vs. 31.4 vs.	21.2 vs. 16.4	4.9 vs.	5.1 vs. 10.0	5.9 vs. 10.0	vs. 5.1 vs.
	66.4 vs. 69.4	48.2 vs. 66.7	30.9 vs. 27.8	vs. 16.7	7.6 vs. 10.9 vs. 0.0	vs. 11.1	vs. 5.6	1.8 vs. 2.8
P-value	0.30	0.013[a]	0.96	0.18	0.08	0.16	0.30	0.70
Histological grade:	7 vs. 237 vs. 5	5 vs. 164 vs.	3 vs. 111 vs.	1 vs. 60 vs. 5	1 vs. 29	1 vs. 23 vs. 2	3 vs. 24 vs. 0	0 vs. 13 vs.
Poor vs. medium		17	10		vs. 4			3
vs. well	58.3 vs. 64.4	41.7 vs. 44.6	25.0 vs. 30.2	8.3 vs. 16.3 vs.	8.3 vs.	8.3 vs. 6.3 vs.	25.0 vs. 6.5	0.0 vs. 3.5
differentiation	vs. 23.8	vs. 81.0	vs. 47.6	23.8	7.9 vs. 19.0	9.5	vs. 0.0	vs. 14.3
P-value	<0.001[a]	0.004[a]	0.22	0.51	0.16	0.49	0.04[a]	0.06
Lymph node metastasis:	104 vs.	83 vs.	50 vs.	26 vs.	16 vs.	16 vs.	14 vs.	3 vs.
Yes vs. no	148	110	76	40	18	12	13	13
	65.0 vs. 58.3	51.9 vs. 43.3	31.3 vs. 29.9	16.3 vs. 15.7	10.0 vs. 7.1	10.0 vs. 4.7	8.8 vs. 5.1	1.9 vs. 5.1
P-value	0.21	0.11	0.86	1	0.39	0.06	0.21	0.16

[a] Statistically significant. MSI-H, high microsatellite instability; MSI-L, low microsatellite instability; MSS, microsatellite stability.

that *TP53* mutations were associated with MSI-L/MSS, location, and histological grade. As a crucial cell-cycle checkpoint and tumor suppressor gene, *TP53* gene deficiency can accelerate tumorigenesis and is associated with poorer survival rates in patients [31]. In the present cohort of Chinese patients with CRC, *TP53* mutations most frequently occurred at R273, R282, R248, and R175 positions, which are hotspot codons in the central DNA binding domain [30]. However, there exists heterogeneity across *TP53* mutation positions in terms of tumor invasion and survival, and thus further exploration of *TP53* status could lead to more accurate individual therapeutic decisions [32]. *KRAS* mutations were more frequent in women in our cohort, and were associated with distant metastasis and

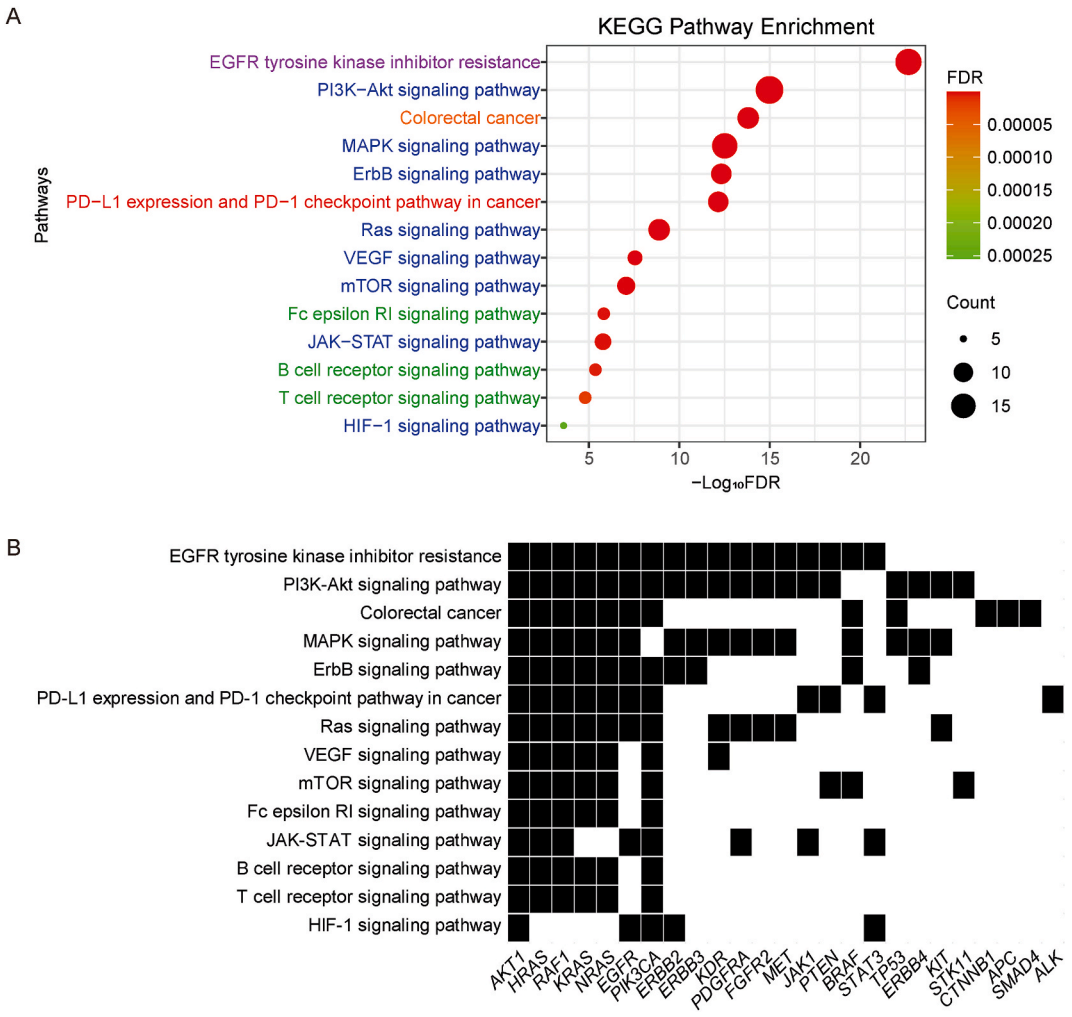

Fig. 3. KEGG pathway enrichment analysis of mutant genes in patients with CRC. (A) Enriched signaling pathway of mutant genes. FDR was calculated by KEGG algorithm. (B) Genes distribution of mutant genes in enriched signaling pathways. KEGG, Kyoto Encyclopedia of Genes and Genomes; FDR, false discovery rate.

malignancy levels, which was consistent with the findings of previous studies [32,33]. *KRAS* plays critical regulatory roles in the downstream stimulation of the EGFR cascade, including the RAS-MAPK and PI3K-Akt signaling pathway, which is involved in CRC tumorigenesis [34]. However, while anti-EGFR monoclonal antibody (mAb) therapy has been widely used in clinical treatment, gene mutations in *RAS* family members significantly impair its efficacy [35]. In the present cohort, 50 % (207/414) of patients carried *RAS* gene mutations. Moreover, in *RAS* wildtype patients, 22.7 % (47/207) of them carried *BRAF*, *PIK3CA* or *PTEN* mutations, suggesting that these patients may not benefit from anti-EGFR mAb therapy [35]. For *APC* gene alterations, despite ranking third in mutation frequency, little correlation was found between *APC* mutation and clinical characteristics in the present study, except for a higher mutation rate in the older population ($P = 0.04$). In addition, 93.7 % (118/126) of patients harboring an *APC* mutation also carried at least one other gene mutation in our panel. The above results are consistent with the fact that *APC* acts as an early and foundational mutation in CRC tumorigenesis [36,37].

In the present study, the MSI status of tumors was tested based on two separate single nucleotide site panels. A total of 26 patients with MSI-H tumors in this cohort all carried at least one detected mutation, which resulted from the diminished capacity of DNA repair mechanisms [38]. In addition, KEGG analysis revealed the enrichment of genes regulating PD-L1 expression and PD-1 checkpoint pathway. For immunotherapy, the use of PD-1/L1 antibody drugs is mainly based on the MSI status of tumors [39,40], and patients with MSI-H type do not benefit from postoperative 5-fluorouracil adjuvant therapy [41–43]. Of note, *BRAF* mutation was significantly associated with MSI state ($P < 0.001$), among which *BRAF* V600E mutations were found in 2.2 % (8/367) of MSI-L/MSS patients and 30.8 % (8/26) of MSI-H patients. However, questions remain about the efficacy of immune checkpoint inhibitor treatment for MSI-H patients with a *BRAF* V600E mutation and metastatic CRC [44]. Thus, genetic mutation detection could provide tumor patients with more accurate therapies.

The NGS approach has shown great potential in assisting the clinical application of individual targeted therapy [45]. In CRC

studies, panels in varying scales are commonly used [46]. NGS is able to sequence a large gene panel at the same time and achieve a higher sensitivity of rare mutations at a relatively low cost [47]. In the present study, a special panel consisting of 73 genes was used; of note, besides the 59 well-known tumor-related hotspot genes, the panel also included 14 genes from the PharmGKB database, which could provide information about how the detected genetic variations may impact drug response. However, due to the recent diagnosis of enrolled participants, we were unable to provide follow-up data on therapy responses so far. At present, the detection of gene mutation and MSI status in patients can help further drug guidance according to the NCCN guidelines for colon or rectal cancer, particularly for patients with metastatic tumors (58 cases in this study), which is the fundamental significance of conducting mutation testing. Nevertheless, our follow-up studies will continue to explore in depth the value of this panel in guiding the individual targeted therapy for patients with CRC.

Our study has some limitations as well. Due to the recent diagnosis of enrolled participants, we were unable to provide follow-up data on recurrence and therapy response. In the future work, we will follow up these CRC patients and analyze the relationship between these mutant genes in this study and the clinical prognosis of CRC patients to guide clinical treatment and drug use.

In conclusion, this study identified genomic alterations in 414 Chinese patients with CRC via the NGS approach using a lab-developed panel and described the correlation between tumor genotype and clinicopathological characteristics. Insights into these oncogenic genetic alterations may help improve individual targeted therapies in the future.

Funding sources

No funding was received.

Ethical statement

All the patients enrolled have signed the written informed consent. This research was approved by the Ethics Committee of Chinese People's Liberation Army Rocket Force Characteristic Medical Center (Approval no.KY2023038) and conducted according to the principles expressed in the Declaration of Helsinki.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Bo Cheng: Writing – original draft, Conceptualization. **Lin Xu:** Writing – original draft, Investigation. **Yunzhi Zhang:** Supervision, Resources. **Huimin Yang:** Resources, Methodology. **Shan Liu:** Investigation. **Shanshan Ding:** Investigation. **Huan Zhao:** Methodology. **Yi Sui:** Writing – review & editing, Visualization, Validation. **Chan Wang:** Writing – review & editing, Validation. **Lanju Quan:** Resources. **Jinhong Liu:** Resources. **Ye Liu:** Methodology. **Hongming Wang:** Visualization, Formal analysis, Data curation. **Zhaoqing Zheng:** Validation. **Xizhao Wu:** Validation. **Jing Guo:** Validation. **Zhaohong Wen:** Project administration. **Ruya Zhang:** Methodology. **Fei Wang:** Supervision. **Hongmei Liu:** Writing – review & editing, Visualization, Formal analysis, Conceptualization. **Suozhu Sun:** Resources, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: HY, HL, YS, CW, HW, ZZ, ZW and YZ reported being employees of Singlera Genomics (Shanghai) Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e29299.

References

[1] F. Ciardiello, D. Ciardiello, G. Martini, et al., Clinical management of metastatic colorectal cancer in the era of precision medicine, CA Cancer J Clin 72 (4) (2022) 372–401, https://doi.org/10.3322/caac.21728.

[2] F. Yagin, A. Alkhateeb, C. Colak, et al., A Fecal-Microbial-Extracellular-Vesicles-based Metabolomics Machine Learning Framework and biomarker discovery for predicting colorectal cancer patients, Metabolites 13 (5) (2023), https://doi.org/10.3390/metabo13050589.

[3] Y. Zhuang, H. Wang, D. Jiang, et al., Multi gene mutation signatures in colorectal cancer patients: predict for the diagnosis, pathological classification, staging and prognosis, BMC Cancer 21 (1) (2021) 380, https://doi.org/10.1186/s12885-021-08108-9.

[4] M. Aiello, N. Vella, C. Cannavo, et al., Role of genetic polymorphisms and mutations in colorectal cancer therapy (Review), Mol. Med. Rep. 4 (2) (2011) 203–208, https://doi.org/10.3892/mmr.2010.408.

[5] J.L. Westra, M. Schaapveld, H. Hollema, et al., Determination of TP53 mutation is more relevant than microsatellite instability status for the prediction of disease-free survival in adjuvant-treated stage III colon cancer patients, J. Clin. Oncol. 23 (24) (2005) 5635–5643, https://doi.org/10.1200/jco.2005.04.096.

[6] G.L. De' Angelis, L. Bottarelli, C. Azzoni, et al., Microsatellite instability in colorectal cancer, Acta Biomed. 89 (9-S) (2018) 97–101, https://doi.org/10.23750/abm.v89i9-S.7960.

[7] C. Luchini, F. Bibeau, M.J.L. Ligtenberg, et al., ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach, Ann. Oncol. 30 (8) (2019) 1232–1243, https://doi.org/10.1093/annonc/mdz116.

[8] J. Taieb, M. Svrcek, R. Cohen, et al., Deficient mismatch repair/microsatellite unstable colorectal cancer: diagnosis, prognosis and treatment, Eur. J. Cancer 175 (2022) 136–157, https://doi.org/10.1016/j.ejca.2022.07.020.

[9] T.M. Pawlik, C.P. Raut, M.A. Rodriguez-Bigas, Colorectal carcinogenesis: MSI-H versus MSI-L, Dis. Markers 20 (4–5) (2004) 199–206, https://doi.org/10.1155/2004/368680.

[10] W. Hou, C. Yi, H. Zhu, Predictive biomarkers of colon cancer immunotherapy: present and future, Front. Immunol. 13 (2022) 1032314, https://doi.org/10.3389/fimmu.2022.1032314.

[11] X. Zhang, T. Wu, X. Cai, et al., Neoadjuvant immunotherapy for MSI-H/dMMR Locally Advanced colorectal cancer: New Strategies and Unveiled Opportunities, Front. Immunol. 13 (2022) 795972, https://doi.org/10.3389/fimmu.2022.795972.

[12] J. Shia, N. Schultz, D. Kuk, et al., Morphological characterization of colorectal cancers in the Cancer Genome Atlas reveals distinct morphology-molecular associations: clinical and biological implications, Mod. Pathol. 30 (4) (2017) 599–609, https://doi.org/10.1038/modpathol.2016.198.

[13] Y. Umeda, T. Nagasaka, Y. Mori, et al., Poor prognosis of KRAS or BRAF mutant colorectal liver metastasis without microsatellite instability, J Hepatobiliary Pancreat Sci 20 (2) (2013) 223–233, https://doi.org/10.1007/s00534-012-0531-9.

[14] K. Fujiyoshi, G. Yamamoto, A. Takahashi, et al., High concordance rate of KRAS/BRAF mutations and MSI-H between primary colorectal cancer and corresponding metastases, Oncol. Rep. 37 (2) (2017) 785–792, https://doi.org/10.3892/or.2016.5323.

[15] D. Huang, W. Sun, Y. Zhou, et al., Mutations of key driver genes in colorectal cancer progression and metastasis, Cancer Metastasis Rev. 37 (1) (2018) 173–187, https://doi.org/10.1007/s10555-017-9726-5.

[16] A. S, A. Chakraborty, S. Patnaik, Clonal evolution and expansion associated with therapy resistance and relapse of colorectal cancer, Mutat. Res. Rev. Mutat. Res. 790 (2022) 108445, https://doi.org/10.1016/j.mrrev.2022.108445.

[17] K. Tong, O.A. Kothari, K.S. Haro, et al., SMAD4 is critical in suppression of BRAF-V600E serrated tumorigenesis, Oncogene 40 (41) (2021) 6034–6048, https://doi.org/10.1038/s41388-021-01997-x.

[18] B. Vogelstein, E.R. Fearon, S.R. Hamilton, et al., Genetic alterations during colorectal-tumor development, N. Engl. J. Med. 319 (9) (1988) 525–532, https://doi.org/10.1056/NEJM198809013190901.

[19] A. Belkadi, A. Bolze, Y. Itan, et al., Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants, Proc Natl Acad Sci U S A 112 (17) (2015) 5473–5478, https://doi.org/10.1073/pnas.1418631112.

[20] S. Kumar, S.K. Razzaq, A.D. Vo, et al., Identifying fusion transcripts using next generation sequencing, Wiley Interdiscip Rev RNA 7 (6) (2016) 811–823, https://doi.org/10.1002/wrna.1382.

[21] L. Tattini, R. D'Aurizio, A. Magi, Detection of genomic structural variants from next-generation sequencing data, Front. Bioeng. Biotechnol. 3 (2015) 92, https://doi.org/10.3389/fbioe.2015.00092.

[22] X. Wang, X. Li, Y. Cheng, et al., Copy number alterations detected by whole-exome and whole-genome sequencing of esophageal adenocarcinoma, Hum Genomics 9 (1) (2015) 22, https://doi.org/10.1186/s40246-015-0044-0.

[23] R. Jiang, B. Zhang, X. Teng, et al., Validating a targeted next-generation sequencing assay and profiling somatic variants in Chinese non-small cell lung cancer patients, Sci. Rep. 10 (1) (2020), https://doi.org/10.1038/s41598-020-58819-5.

[24] Y. Wang, H. Liu, Y. Hou, et al., Performance validation of an amplicon-based targeted next-generation sequencing assay and mutation profiling of 648 Chinese colorectal cancer patients, Virchows Arch. 472 (6) (2018) 959–968, https://doi.org/10.1007/s00428-018-2359-4.

[25] World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects, JAMA 310 (20) (2013) 2191–2194, https://doi.org/10.1001/jama.2013.281053.

[26] C.R. Boland, S.N. Thibodeau, S.R. Hamilton, et al., A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer, Cancer Res. 58 (22) (1998) 5248–5257.

[27] E. Fountzilas, V. Kotoula, G.A. Koliou, et al., Pathogenic mutations and overall survival in 3,084 patients with cancer: the Hellenic Cooperative oncology group precision medicine Initiative, Oncotarget 11 (1) (2020) 1–14, https://doi.org/10.18632/oncotarget.27338.

[28] A.L. Russo, D.R. Borger, J. Szymonifka, et al., Mutational analysis and clinical correlation of metastatic colorectal cancer, Cancer 120 (10) (2014) 1482–1490, https://doi.org/10.1002/cncr.28599.

[29] Y.Y. Chang, P.C. Lin, H.H. Lin, et al., Mutation spectra of RAS gene family in colorectal cancer, Am. J. Surg. 212 (3) (2016) 537–544.e3, https://doi.org/10.1016/j.amjsurg.2016.02.013.

[30] M.C. Liebl, T.G. Hofmann, The role of p53 signaling in colorectal cancer, Cancers 13 (9) (2021), https://doi.org/10.3390/cancers13092125.

[31] A.J. Munro, S. Lain, D.P. Lane, P53 abnormalities and outcomes in colorectal cancer: a systematic review, Br. J. Cancer 92 (3) (2005) 434–444, https://doi.org/10.1038/sj.bjc.6602358.

[32] O. Hassin, N.B. Nataraj, M. Shreberk-Shaked, et al., Different hotspot p53 mutants exert distinct phenotypes and predict outcome of colorectal cancer patients, Nat. Commun. 13 (1) (2022) 2800, https://doi.org/10.1038/s41467-022-30481-7.

[33] M. Meng, K. Zhong, T. Jiang, et al., The current understanding on the impact of KRAS on colorectal cancer, Biomed. Pharmacother. 140 (2021) 111717, https://doi.org/10.1016/j.biopha.2021.111717.

[34] L. Barault, N. Veyrie, V. Jooste, et al., Mutations in the RAS-MAPK, PI(3)K (phosphatidylinositol-3-OH kinase) signaling network correlate with poor survival in a population-based series of colon cancers, Int. J. Cancer 122 (10) (2008) 2255–2259, https://doi.org/10.1002/ijc.23388.

[35] W. De Roock, V. De Vriendt, N. Normanno, et al., KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer, Lancet Oncol. 12 (6) (2011) 594–603, https://doi.org/10.1016/s1470-2045(10)70209-6.

[36] S. Narayan, D. Roy, Role of APC and DNA mismatch repair genes in the development of colorectal cancers, Mol. Cancer 2 (2003) 41, https://doi.org/10.1186/1476-4598-2-41.

[37] L. Zhang, J.W. Shay, Multiple roles of APC and its therapeutic implications in colorectal cancer, J Natl Cancer Inst 109 (8) (2017), https://doi.org/10.1093/jnci/djw332.

[38] C.R. Boland, A. Goel, Microsatellite instability in colorectal cancer, Gastroenterology 138 (6) (2010) 2073–2087.e3, https://doi.org/10.1053/j.gastro.2009.12.064.

[39] D.T. Le, J.N. Durham, K.N. Smith, et al., Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade, Science 357 (6349) (2017) 409–413, https://doi.org/10.1126/science.aan6733.

[40] M.J. Overman, S. Lonardi, K.Y.M. Wong, et al., Durable clinical benefit with Nivolumab Plus Ipilimumab in DNA mismatch repair-deficient/microsatellite instability-high metastatic colorectal cancer, J. Clin. Oncol. 36 (8) (2018) 773–779, https://doi.org/10.1200/jco.2017.76.9901.

[41] J.M. Carethers, E.J. Smith, C.A. Behling, et al., Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer, Gastroenterology 126 (2) (2004) 394–401, https://doi.org/10.1053/j.gastro.2003.12.023.

[42] R. Jover, P. Zapater, A. Castells, et al., Mismatch repair status in the prediction of benefit from adjuvant fluorouracil chemotherapy in colorectal cancer, Gut 55 (6) (2006) 848–855, https://doi.org/10.1136/gut.2005.073015.

[43] C.M. Ribic, D.J. Sargent, M.J. Moore, et al., Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer, N. Engl. J. Med. 349 (3) (2003) 247–257, https://doi.org/10.1056/NEJMoa022289.
[44] G. Mauri, E. Bonazzina, A. Amatu, et al., The Evolutionary landscape of treatment for BRAF(V600E) mutant metastatic colorectal cancer, Cancers 13 (1) (2021), https://doi.org/10.3390/cancers13010137.
[45] D. Qin, Next-generation sequencing and its clinical application, Cancer Biol Med 16 (1) (2019) 4–10, https://doi.org/10.20892/j.issn.2095-3941.2018.0055.
[46] F. Del Vecchio, V. Mastroiaco, A. Di Marco, et al., Next-generation sequencing: recent applications to the analysis of colorectal cancer, J. Transl. Med. 15 (1) (2017) 246, https://doi.org/10.1186/s12967-017-1353-y.
[47] Y. Zhong, F. Xu, J. Wu, et al., Application of next generation sequencing in laboratory medicine, Ann Lab Med 41 (1) (2021) 25–43, https://doi.org/10.3343/alm.2021.41.1.25.