



Article

# High Concordance of Genomic Profiles between Primary and Metastatic Colorectal Cancer

Seung Eun Lee <sup>1</sup>, Ha Young Park <sup>2,\*</sup>, Dae-Yong Hwang <sup>3</sup> and Hye Seung Han <sup>1,\*</sup>

<sup>1</sup> Department of Pathology, School of Medicine, Konkuk University, Seoul 05030, Korea; 20150063@kuh.ac.kr

<sup>2</sup> Department of Pathology, Busan Paik Hospital, College of Medicine, Inje University, Busan 47392, Korea

<sup>3</sup> Department of Surgery, School of Medicine, Konkuk University, Seoul 05030, Korea; hwangcrc@kuh.ac.kr

\* Correspondence: pmint00@naver.com (H.Y.P.); 20040002@kuh.ac.kr (H.S.H.)

**Abstract:** The comparison of the genetic profiles between primary and metastatic colorectal cancer (CRC) is needed to enable the discovery of useful therapeutic targets against metastatic CRCs. We performed the targeted next generation sequencing assay of 170 cancer-associated genes for 142 metastatic CRCs, including 95 pairs of primary and metastatic CRCs, to reveal their genomic characteristics and to assess the genetic heterogeneity. The most frequently mutated gene in primary and metastatic CRCs was *APC* (71% vs. 65%), *TP53* (54% vs. 57%), *KRAS* (45% vs. 44%), *PIK3CA* (16% vs. 19%), *SMAD4* (15% vs. 14%) and *FBXW7* (11% vs. 11%). The concordance in the top six frequently mutated genes was 85%, on average. The overall mutation frequencies were consistent with two sets of public data (TCGA and MSKCC). To the author's knowledge, this is the first study to compare the genetic profiles of our cohort with that of the metastatic CRCs from MSKCC. Comparative sequencing analysis between primary and metastatic CRCs revealed a high degree of genetic concordance in the current clinically actionable genes. Therefore, the genetic investigation of archived primary tumor samples with the challenges of obtaining an adequate sample from metastatic sites appears to be sufficient for the application of cancer precision medicine in the metastatic setting.

**Keywords:** metastatic colorectal cancer; genomic profiles; high concordance between primary and metastatic colorectal cancer



**Citation:** Lee, S.E.; Park, H.Y.; Hwang, D.-Y.; Han, H.S. High Concordance of Genomic Profiles between Primary and Metastatic Colorectal Cancer. *Int. J. Mol. Sci.* **2021**, *22*, 5561. <https://doi.org/10.3390/ijms22115561>

Academic Editor: Carlos Flores

Received: 16 April 2021

Accepted: 16 May 2021

Published: 24 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Colorectal cancer (CRC) is one of the well-established tumor types to be considered as a genetic disease in which the multiple and sequential accumulation of genetic alterations underlies the development and progression to carcinoma and metastasis. The inactivation of *APC* mutations, activation of *KRAS* mutations, and the diverse mutations in *TP53*, *PIK3CA*, and *SMAD4*, TGF- $\beta$  pathway genes, drive the development and evolution of a malignant CRC [1].

A comprehensive investigation of the genomic landscape of the early stages of CRC, was reported by The Cancer Genome Atlas (TCGA) Network Network [2]. More recently, metastatic CRCs were also analyzed by using MSK-IMPACT, a capture-based next generation sequencing (NGS) platform [3]. Several studies have performed the analysis of the comparative genetic sequencing of paired primary and metastatic CRC [4–7]. The majority have shown a high degree of concordance in the genetic profile between primary and metastatic CRC [4–6]. Especially, early recurrent genetic alterations such as *APC*, *KRAS*, *NRAS*, and *BRAF*, involving colorectal carcinogenesis, were highly concordant in matched pairs of primary and metastatic CRC [5,6].

Genetic intratumor heterogeneity within tumors often occurs as a result of the progressive accumulation of genetic alterations during the spatial and temporal evolution of the tumor. More recently, several studies have reported that advanced CRCs harbor extensive intratumor heterogeneity, shaped by neutral evolution during tumor evolution [8,9]. Furthermore, Saito et al. demonstrated that the evolutionary principle shaping

genetic intratumor heterogeneity shifts from Darwinian to neutral evolution during CRC progression [9].

The comparison of the genetic profile between primary and metastatic CRCs is needed to enable the discovery of useful therapeutic targets against metastatic CRCs. In this study, we performed the targeted next generation sequencing (NGS) assay of 170 cancer-associated genes for 142 metastatic CRCs, including 95 pairs of primary and metastatic colorectal samples, to reveal their genomic characteristics and to assess the genetic heterogeneity between primary and metastatic CRCs.

## 2. Results

### 2.1. Clinicopathologic Characteristics of Patients

A total of 146 patients were analyzed in this study. From 95 patients, pairs of primary and metastatic colon cancer samples were analyzed to compare the genetic profiles. An additional 4 primary only and 47 metastatic only samples were also analyzed. The clinicopathologic characteristics of 142 metastatic tumors (including 95 paired with primary tumor and 47 singleton) are summarized in Table 1. The median age at diagnosis was 61 years (range, 34–89). The most common sites of metastasis were the liver (54.2%) and the lung (24.7%), followed by the abdominopelvic cavity (12.7%) and central nervous system and soft and bone metastases were rare. In total, 78 (54.9%) and 64 (45.1%) of 142 patients developed synchronous and metachronous metastasis, respectively. Only 4.9% of metastatic CRCs displayed an MSI-H genotype/phenotype.

**Table 1.** Clinicopathologic characteristics of 142 metastatic tumors.

Variable	Category	Total (n = 142)	
Age	Mean ± SD	60.51 ± 11.17	
	Median	61	
	Range (min–max)	34–89	
Gender	Male	95	66.90%
	Female	47	33.10%
Site of metastasis	Liver	77	54.23%
	Lung	35	24.65%
	Abdominopelvic cavity	18	12.68%
	Soft tissue	6	4.23%
	Brain	2	1.41%
	Lymph node	2	1.41%
	Bone	1	0.70%
	Skin	1	0.70%
Location of primary cancer	Ascending colon	16	11.30%
	Hepatic flexure	6	4.20%
	Transverse colon	5	3.50%
	Splenic flexure	2	1.40%
	Descending colon	6	4.20%
	Rectosigmoid	107	75.40%
Patterns of metastasis	Synchronous	78	54.90%
	Metachronous	64	45.10%
T stage	1	3	2.11%
	2	7	4.93%
	3	101	71.13%
	4	26	18.31%
	NA	5	3.52%

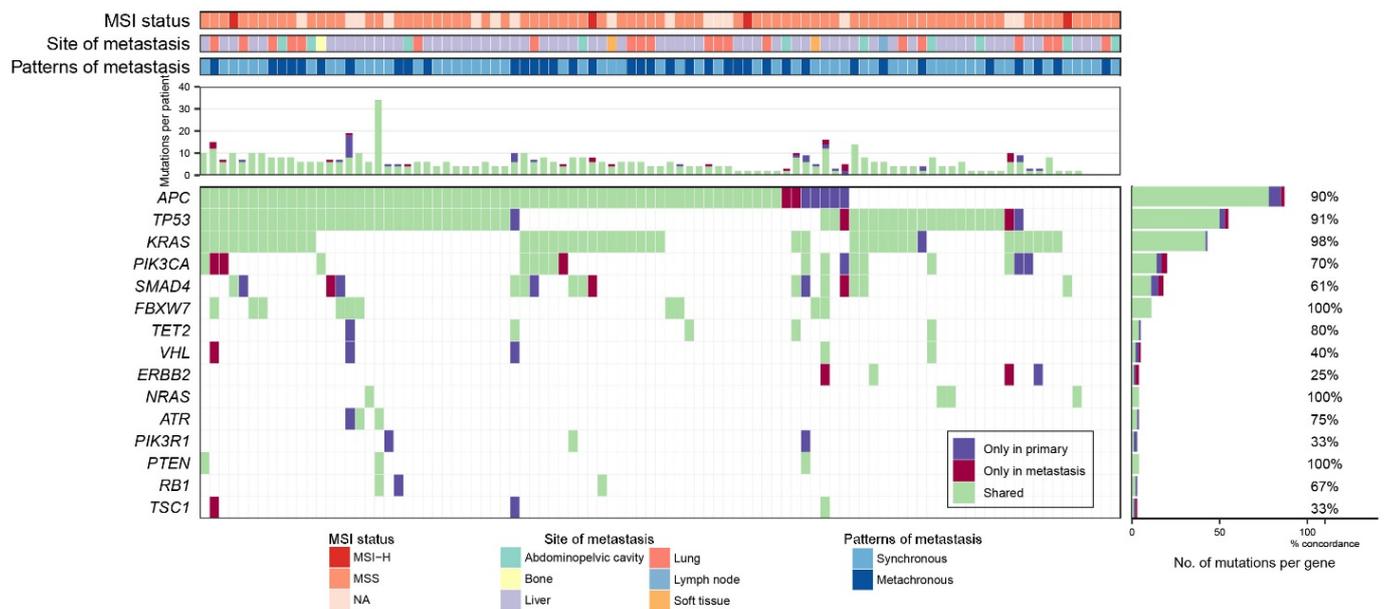
Table 1. Cont.

Variable	Category	Total (n = 142)	
N stage	0	39	27.46%
	1	48	33.80%
	2	48	33.80%
	NA	7	4.93%
Grade	1	2	1.41%
	2	132	92.96%
	3	4	2.82%
	NA	4	2.82%
Vascular invasion	Absent	117	82.39%
	Present	20	14.08%
	NA	5	3.52%
Lymphatic invasion	Absent	72	50.70%
	Present	65	45.77%
	NA	5	3.52%
Perineural invasion	Absent	97	68.31%
	Present	40	28.17%
	NA	5	3.52%
MSI status	MSI-H	7	4.93%
	MSS	107	75.35%
	NA	21	14.79%

## 2.2. Genomic Profiling of 95 Paired Samples

After the filtering steps to identify the clinically significant mutations described in the Method section, a total of 318 mutations of 51 genes, including 243 missense and nonsense single nucleotide variants (SNVs), 52 insertion and deletions (indels) and 23 splice site mutations were identified in 95 primary-tumor pairs. The full list of variants can be found in Table S1. Variants occurred in more than 3 out of 95 pairs and are summarized in Figure 1. Of the 318 mutations, 81% (258/318) were found in both primary and metastatic CRCs, 12% (37/318) were found only in primary CRCs, and 7% (23/318) were found only in metastatic CRCs. Concordance was different according to the type of variant, 85% for SNVs, 73% for indels and 57% for splice site mutations. The concordance of the variants in the top six genes was 85% on average, the lowest in SMAD4 was 61% and the highest in FBXW7 was 100%.

The average number of concordant variants per sample was 5.4 and discordant variants per sample was 0.6. The discordant variants were caused more frequently by primary specific variants (37/60, 62%), compared with the metastatic-specific variants (23/60, 38%). The portion of concordant mutations was high in FBXW7, NRAS, PTEN, and BRCA2 (100% concordance) gene, whereas the portion of discordant mutations was high in ERBB2, PIK3R1, TSC1, and VHL (concordance  $\leq 40\%$ ) genes. The mean concordance of variants in all 95 paired primary and metastatic CRCs was 87%. Concordance was significantly higher in the synchronous group than in the metachronous group (92% vs. 86%,  $p = 0.0124$ ). The ratio of primary specific variants was significantly lower in the synchronous group than in the metachronous group (4% vs. 11%,  $p = 0.0016$ ).



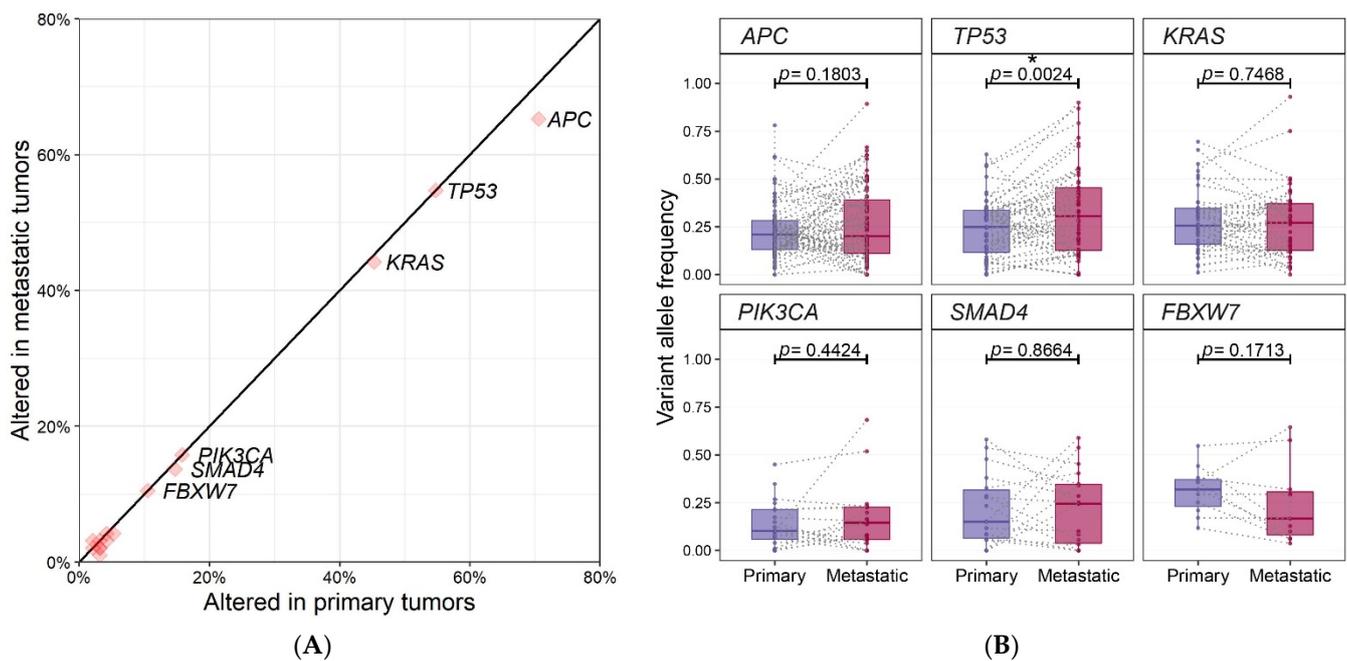
**Figure 1.** Overview of mutational profile of 95 pairs of primary-metastatic colorectal cancers. Each column represents a patient. The top 4 panels show MSI status, site of metastasis, patterns of metastasis and number of variant counts per patients. The bottom panel show mutational status, shared, only in primary and only in metastasis. When there are 2 or more mutation type presents, (i.e., ‘shared’ and ‘primary only’) ‘shared’ is preferentially used. The right panel represents the number of patients with variants as bar graph and concordance of the variants as percentage of corresponding genes.

The most frequently mutated gene in primary CRCs was *APC* (71%, 67/95), followed by *TP53* (54%, 52/95), *KRAS* (45%, 43/95), *PIK3CA* (16%, 15/95), *SMAD4* (15%, 14/95) and *FBXW7* (11%, 10/95). In metastatic CRCs, the order was the same as the primary group, *APC* (65%, 62/95), followed by *TP53* (57%, 54/95), *KRAS* (44%, 42/95), *PIK3CA* (19%, 18/95), *SMAD4* (14%, 13/95) and *FBXW7* (11%, 10/95). The frequency of the top 6 gene mutations was almost the same in primary and metastatic CRCs (Figure 2A). The frequency of mutations was numerically higher in primary tumors than in metastatic tumors in *APC*, although, there was no overall statistical difference in the frequency of mutations between primary and metastatic tumors.

We also compared the variant allele frequency (VAF) of the frequently mutated top six genes between primary and metastatic CRCs. (Figure 2B and Table 2). Of the six genes, VAFs of *TP53* mutations in metastatic CRCs were higher than those of primary CRCs ( $0.24\% \pm 0.15\%$  in primary CRCs vs.  $0.33\% \pm 0.22\%$  in metastatic CRCs), which was statistically significant by paired two sample t-test ( $p = 0.0024$ ).

**Table 2.** Variant allele frequency (VAF) of frequently mutated top six genes between primary and metastatic CRCs.

Genes	No. of variants	VAF of Primary Tumors (Mean $\pm$ SD)	VAF of Metastatic Tumors (Mean $\pm$ SD)	<i>p</i> -Value
<i>APC</i>	87	0.23 $\pm$ 0.14	0.26 $\pm$ 0.19	0.1803
<i>TP53</i>	55	0.24 $\pm$ 0.15	0.33 $\pm$ 0.22	0.0024
<i>KRAS</i>	43	0.28 $\pm$ 0.16	0.27 $\pm$ 0.19	0.7468
<i>PIK3CA</i>	20	0.14 $\pm$ 0.12	0.17 $\pm$ 0.17	0.4424
<i>SMAD4</i>	18	0.21 $\pm$ 0.19	0.22 $\pm$ 0.20	0.8664
<i>FBXW7</i>	11	0.31 $\pm$ 0.12	0.24 $\pm$ 0.21	0.1713



**Figure 2.** (A) Comparison of mutational frequencies between 95 pairs of primary and metastatic tumors. It is the proportion of patients with mutations in the gene. Top six genes are annotated. (B) Comparison of variant allele frequencies of top six genes between 95 pairs of primary and metastatic tumors. Boxplots show the distribution of VAFs in each group and dots connected by dashed-lines represent same patients. Statistical significance is based on paired two sample *t*-test and  $p < 0.05$  is marked with asterisk (\*).

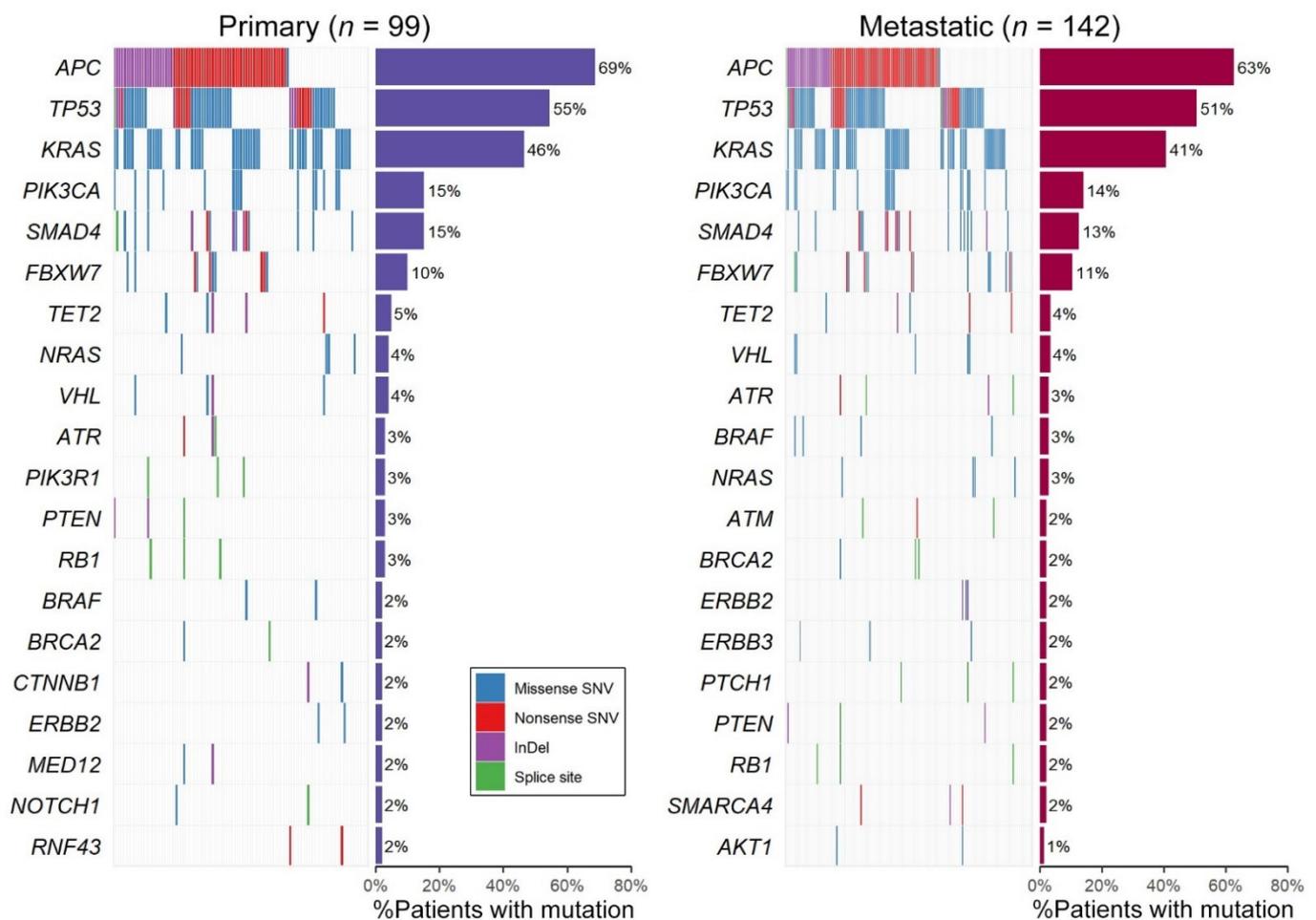
No discordant mutation in *FBXW7* was observed and most of the VAFs in metastatic CRCs were lower than those of primary CRCs (9/11), which is tumor cells harboring a *FBXW7* mutation might be subclonal in metastatic CRCs. No obvious differences in the VAFs of *APC*, *KRAS* and *PIK3CA* were detected between primary and metastatic CRCs.

Mutational counts per sample did not show significant differences according to the patterns of metastasis (synchronous vs. metachronous), location of metastasis (liver vs. others) and MSI status.

### 2.3. Genomic Profiling of All 142 Samples Including Primary/Metastatic Singletons

When analyzing all samples, including unpaired primary or metastatic only in each primary and metastatic group, the top six genes were the same in both groups (Figure 3). The variants in *APC* were mostly the truncating type, 72% nonsense SNVs, 26% indels and 2% splice site mutations in the metastatic group. For *TP53*, 66% were missense SNVs encoding p.R175H (16%), p.R248Q/W (8%) or p.R273H/S (5%) altered proteins and 34% were the truncating type including 22% nonsense, 8% indels and 4% splice site mutations in metastatic group. All of the variants of the two oncogenes, *KRAS* and *PIK3CA* occurred in mutational hotspots. The mutational frequencies of the genes, except for the top six genes, were less than 5% and the rank of the mutational frequencies showed no significant difference. The genes with variants exclusively in the metastatic group were *AKT1*, *CDH1*, *BAP1*, *DNMT3A*, *FOXL2*, and *MSH6*. These genes are probable tumor suppressor genes according to the Cancer Gene Census [10], except for *AKT1*.

The clinicopathological factors showed statistically significant differences by mutational status. (Table 3) MSI-H ratio was higher in the *SMAD4* mutated group. The frequency of *TP53* mutations in synchronous metastasis were higher than those of metachronous metastasis. Overall survival was not different according to mutational status of the top six genes. Recurrence occurred significantly earlier in patients with *FBXW7* mutant (137 days in *FBXW7* mutant vs. 294 days in wild type,  $p = 0.0221$ ).



**Figure 3.** Landscape of mutations in 99 primary and 142 metastatic groups (including 95 pairs of primary and metastatic tumors). Each column represents a single patient. The percentage of mutations across each group is represented by vertical histograms.

**Table 3.** Association between mutational status and clinicopathological features.

Mutational Status	MSI Status				p Value	Location of Primary Cancer				p Value	Patterns of Metastasis			p Value
	Total Patients (n = 118)	MSI-H (n = 7)	MSS (n = 111)			Total Patients (n = 142)	Liver (n = 77)	Others (n = 65)			Total Patients (n = 142)	Synchronous (n = 78)	Metachronous (n = 64)	
APC														
Wild type (%)	47 (39.83)	3 (42.86)	44 (39.64)		53 (37.32)	24 (31.17)	29 (44.62)		53 (37.32)	32 (41.03)	21 (32.81)		0.41	
Mutated (%)	71 (60.17)	4 (57.14)	67 (60.36)	>0.99	89 (62.68)	53 (68.83)	36 (55.38)	0.14	89 (62.68)	46 (58.97)	43 (67.19)			
TP53														
Wild type (%)	58 (49.15)	5 (71.43)	53 (47.75)		70 (49.3)	33 (42.86)	37 (56.92)		70 (49.3)	32 (41.03)	38 (59.38)		0.04	
Mutated (%)	60 (50.85)	2 (28.57)	58 (52.25)	0.27	72 (50.7)	44 (57.14)	28 (43.08)	0.13	72 (50.7)	46 (58.97)	26 (40.62)			
KRAS														
Wild type (%)	69 (58.47)	3 (42.86)	66 (59.46)		84 (59.15)	49 (63.64)	35 (53.85)		84 (59.15)	52 (66.67)	32 (50)		0.07	
Mutated (%)	49 (41.53)	4 (57.14)	45 (40.54)	0.45	58 (40.85)	28 (36.36)	30 (46.15)	0.31	58 (40.85)	26 (33.33)	32 (50)			
PIK3CA														
Wild type (%)	99 (83.9)	6 (85.71)	93 (83.78)		122 (85.92)	64 (83.12)	58 (89.23)		122 (85.92)	68 (87.18)	54 (84.38)		0.81	
Mutated (%)	19 (16.1)	1 (14.29)	18 (16.22)	>0.99	20 (14.08)	13 (16.88)	7 (10.77)	0.42	20 (14.08)	10 (12.82)	10 (15.62)			
SMAD4														
Wild type (%)	102 (86.44)	3 (42.86)	99 (89.19)		124 (87.32)	66 (85.71)	58 (89.23)		124 (87.32)	67 (85.9)	57 (89.06)		0.76	
Mutated (%)	16 (13.56)	4 (57.14)	12 (10.81)	0.01	18 (12.68)	11 (14.29)	7 (10.77)	0.71	18 (12.68)	11 (14.1)	7 (10.94)			
FBXW7														
Wild type (%)	106 (89.83)	7 (100)	99 (89.19)		127 (89.44)	66 (85.71)	61 (93.85)		127 (89.44)	68 (87.18)	59 (92.19)		0.49	
Mutated (%)	12 (10.17)	0 (0)	12 (10.81)	>0.99	15 (10.56)	11 (14.29)	4 (6.15)	0.19	15 (10.56)	10 (12.82)	5 (7.81)			

2.4. Comparing with Public Data: 99 Primary CRCs vs. TCGA and 142 Metastatic CRCs vs. MSKCC

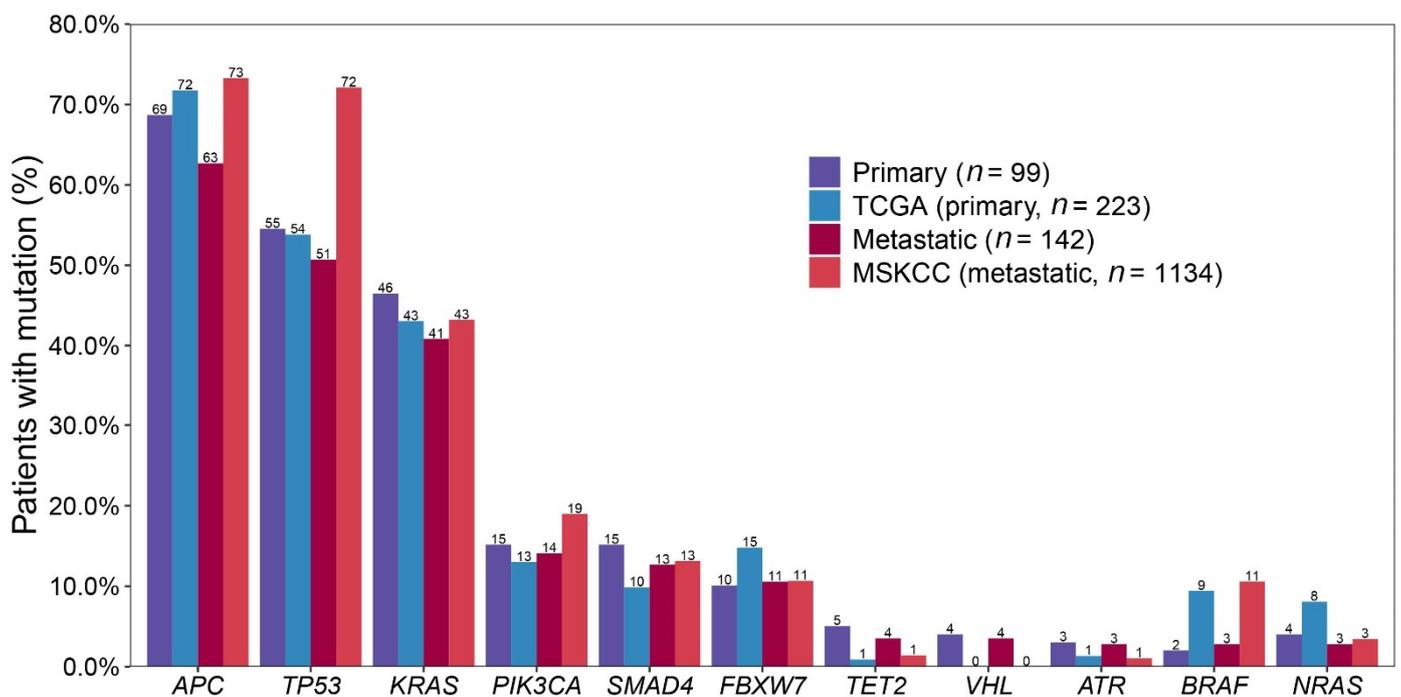
We compared our data with public cancer datasets from cBioPortal. (<https://www.cbioportal.org/datasets>, accessed on 19 May 2021) ‘Colorectal Adenocarcinoma (TCGA, Firehose Legacy)’ (TCGA, n = 223) and ‘Metastatic Colorectal Cancer (MSKCC, Cancer Cell 2018)’ (MSKCC, n = 1134) for primary and metastatic control, respectively. The top 20 genes of each group are listed in Table 4. Of the top 20 genes in TCGA or MSKCC,

*AMER1*, *SOX9*, *ARID1A* and *TCF7L2* were not included in our cancer panel. TCGA, the datasets with genetic profiles of primary CRCs and MSKCC, those of metastatic CRCs were compared with the results in this study (Figure 4).

**Table 4.** Top 20 frequently mutated genes in 4 groups.

Primary (n = 99)				Metastatic (n = 142)				TCGA (n = 223)				MSKCC (n = 1134)			
Order	Gene	# Mutated Patients	% Mutation	Order	Gene	# Mutated Patients	% Mutation	Order	Gene	# Mutated Patients	% Mutation	Order	Gene	# Mutated Patients	% Mutation
1	APC	68	69%	1	APC	89	63%	1	APC	160	72%	1	APC	831	73%
2	TP53	54	55%	2	TP53	72	51%	2	TP53	120	54%	2	TP53	818	72%
3	KRAS	46	46%	3	KRAS	58	41%	3	KRAS	96	43%	3	KRAS	490	43%
4	PIK3CA	15	15%	4	PIK3CA	20	14%	4	FBXW7	33	15%	4	PIK3CA	215	19%
5	SMAD4	15	15%	5	SMAD4	18	13%	5	PIK3CA	29	13%	5	SMAD4	149	13%
6	FBXW7	10	10%	6	FBXW7	15	11%	6	SMAD4	22	10%	6	FBXW7	121	11%
7	TET2	5	5%	7	TET2	5	4%	7	BRAF	21	9%	7	BRAF	120	11%
8	NRAS	4	4%	8	VHL	5	4%	8	AMER1	19	9%	8	SOX9	99	9%
9	VHL	4	4%	9	ATR	4	3%	9	NRAS	18	8%	9	ARID1A	83	7%
10	ATR	3	3%	10	BRAF	4	3%	10	ARID1A	12	5%	10	TCF7L2	77	7%
11	PIK3R1	3	3%	11	NRAS	4	3%	11	ATM	10	4%	11	PTEN	68	6%
12	PTEN	3	3%	12	ATM	3	2%	12	ERBB3	10	4%	12	RNF43	67	6%
13	RBI	3	3%	13	BRCA2	3	2%	13	SOX9	10	4%	13	AMER1	49	4%
14	BRAF	2	2%	14	ERBB2	3	2%	14	SMAD2	9	4%	14	ATM	45	4%
15	BRCA2	2	2%	15	ERBB3	3	2%	15	PIK3R1	8	4%	15	PIK3R1	40	4%
16	CTNNB1	2	2%	16	PTCHI	3	2%	16	CREBBP	7	3%	16	NRAS	39	3%
17	ERBB2	2	2%	17	PTEN	3	2%	17	FBN1	7	3%	17	B2M	37	3%
18	MED12	2	2%	18	RBI	3	2%	18	PTEN	7	3%	18	BRCA2	35	3%
19	NOTCH1	2	2%	19	SMARCA4	3	2%	19	BCOR	6	3%	19	SMAD3	32	3%
20	RNF43	2	2%	20	AKT1	2	1%	20	ELF3	6	3%	20	TGFBR2	30	3%

The top six genes were common in all groups, the order showed differences in TCGA.



**Figure 4.** Most commonly mutated gene frequencies of present study compared with TCGA dataset as control for 99 primary group, and MSKCC dataset as control for 142 metastatic group.

When comparing two public datasets, TCGA and MSKCC, the genes with significantly lower mutational frequency in MSKCC than in TCGA were *AMER1* (9% vs. 4%,  $p = 0.0326$ ) and *NRAS* (8% vs. 3%,  $p = 0.0463$ ). The genes with significantly higher mutational frequency in MSKCC than in TCGA were *TP53* (53% vs. 72%,  $p < 0.0001$ ), *PIK3CA* (13% vs. 19%,  $p = 0.0188$ ), and *SOX9* (4% vs. 9%,  $p = 0.0088$ ). The genes showing significant difference between two public datasets were *NRAS*, *TP53*, and *PIK3CA*, and these genes showed similar mutational frequencies between the primary and metastatic CRCs of the present study.

When comparing the primary group of the present study with TCGA, the genes showing a difference in variant rate  $\geq 5\%$  were *SMAD4* (15% vs. 10%,  $p = 0.1995$ ), *FBXW7* (10% vs. 15%,  $p = 0.2225$ ), and *BRAF* (2% vs. 9%,  $p = 0.0022$ ). Between the metastatic group

of the present study and MSKCC, *APC* (63% vs. 73%,  $p = 0.0129$ ), *TP53* (51% vs. 72%,  $p < 0.0001$ ), *PIK3CA* (14% vs. 19%,  $p = 0.1209$ ), and *BRAF* (3% vs. 11%,  $p < 0.0001$ ) showed differences of more than 5%.

*APC* mutations were mostly the truncated type in MSKCC (1268/1217 mutations, 99.8%) as in our results of the present data. In terms of types of variants, indel ratio was significantly higher in MSKCC than our metastatic group (41% vs. 26%,  $p = 0.0004$ ). Considering *TP53* mutations, ratio of missense SNVs were 66% in both the metastatic group of the present study and MSKCC, and nonsense SNV ratio was higher in the metastatic group of the present study than in MSKCC (22% vs. 14%). However, indel ratio was significantly lower in the metastatic group of the present study than in MSKCC (8% vs. 15%,  $p = 0.0488$ ). *PIK3CA* mutations occurred at p.H1047 were 36% in the metastatic group of the present study and 17% in MSKCC, at p.E545 were 32% vs. 27%. Ninety-five percent of *PIK3CA* variants were found in the amino acid positions of 1047, 542, 545 and 546 in the metastatic group of the present study, 32% of variants were found at other positions in MSKCC. *BRAF* V600E mutation was only found in two patients of the present study.

### 3. Discussion

We performed targeted NGS of 170 cancer-associated genes in 142 metastatic CRCs, including 95 pairs of primary and metastatic CRCs to define the mutational concordance of these genes in primary and metastatic CRCs. Furthermore, we compared our data with the public cancer datasets, TCGA ( $n = 223$ ) [2] and MSKCC ( $n = 1134$ ) [3]. The previous comparative sequencing studies between primary and metastatic CRCs were only compared with the data of TCGA [6,11]. TCGA analyzed only the primary CRCs, the majority of which were derived from the early stages of CRCs [2]. Meanwhile, our study cohort, which consists of all patients with metastatic CRCs, and the MSKCC CRC cohort, which also had more aggressive and advanced CRCs, were distinct from the TCGA cohort [3]. Genomic analysis in the MSKCC thus provided insights into more metastatic CRCs that were not evident in the TCGA CRCs cohort. Therefore, we compared genetic profiles of primary CRCs in our study with TCGA, and those of the metastatic CRCs with the MSKCC in our study, respectively. Our data is significant in comparing primary and metastatic CRCs in pairs, as well as comparing two large public data, TCGA and MSKCC, representative of primary and metastatic CRCs, under the same conditions. To the author's knowledge, this is the first study to compare the genetic profile of our cohort with that of the metastatic CRCs from MSKCC dataset.

We identified that the frequency of recurrent mutations of *APC*, *KRAS*, *PIK3CA*, *FBXW7*, and *SMAD4* were consistent with previous reports on metastatic CRCs [3,12,13]. Overall concordance of the clinically significant mutations between primary and metastatic CRCs was 81%. This concordance rose to 85% for the six most recurrent mutations occurring in CRCs. The six most recurrent mutations were known as the CRC driver genes. These results are consistent with those of previous studies on comparative sequencing [3,5,14]. In MSKCC, a high level of genomic concordance was also identified between the primary and metastatic CRCs [3]. Therefore, a low degree of genetic heterogeneity between primary and metastatic CRCs with respect to driver mutations of CRCs. This supported that the main driver genetic alterations involving colorectal carcinogenesis was maintained during the evolution of tumor metastasis. Among the driver genes, the concordance was high in *KRAS* (98%), *NRAS* (100%), *APC* (90%), which are early or universal mutations in CRCs. On the other hand, the concordance was relatively low in *PIK3CA* (70%), *SMAD4* (61%), which are known as later mutations. This observation was consistent with prior studies [11] and could be explained through heterogeneous clonal evolution [15].

Especially, our result showed that with *KRAS* (98%), *NRAS* (100%), there was very high concordance between primary and metastatic CRCs. A meta-analysis of all published studies between 1991 and 2018 reporting on biomarker concordance between primary and metastatic CRCs, a very high median biomarker concordance for *KRAS* (93%), *NRAS* (100%), *BRAF* (99.4%), *PIK3CA* (93%) was reported, whereas meta-analytic pooled discordance

was 8% for *KRAS*, 8% for *BRAF*, and 7% for *PIK3CA* in 61 studies, including 3565 patient samples [16]. *KRAS* and *NRAS* are OncoKB level-1 resistance biomarkers for anti-EGFR (epidermal growth factor receptor) antibodies (cetuximab and panitumumab), mutational testing of these genes has now been incorporated into the National Comprehensive Cancer Network (NCCN) guidelines for the treatment of patients with metastatic CRCs [17]. These results are also consistent with the recommendation that molecular analysis of the primary tumor is representative of the genetic characteristic of the metastatic tumor [17]. Therefore, the genetic investigation of archived primary tumor samples with challenges in getting adequate samples from metastatic sites appears to be sufficient for the application of cancer precision medicine in the metastatic setting.

Concordance between primary and metastatic CRCs was significantly higher in the synchronous group than in the metachronous group. Synchronous metastatic tumors were mainly treatment-naïve while the majority of patients with metachronous metastatic tumors have received systemic treatment such as chemotherapy for the primary tumor, which may influence the mutational concordance between the primary and the subsequent tumor.

The genes with significantly lower mutational frequency in the metastatic group of the present study than in the MSKCC were *TP53*, and *APC*, whereas genes showing significant difference of mutational frequency between the primary group of the present study and TCGA were not identified. Considering the distribution of the types of variants, the MSKCC group showed a significantly higher proportion of indels among the total variants than those of the present study and TCGA. (35% for MSKCC, 14% for present study and TCGA,  $p < 0.0001$ ). *TP53* and *APC* are the most representative tumor suppressor genes which commonly suffer loss of their function by deletion mutations. In the MSKCC group, due to a more efficient indel calling algorithm, more indels may have been detected than in the present study and made a difference in the mutational frequencies in the two genes commonly altered by deletion mutation.

*TP53* alterations in the MSKCC group, the genetic profiles of metastatic CRCs, were the only genomic alteration significantly enriched in metastatic CRCs, which shows *TP53* alterations are selectively enriched in metastatic CRCs [3]. In our study, VAFs of *TP53* mutations in metastatic CRCs were significantly higher than those of the primary CRCs, although there was no overall statistical difference in the frequency of mutations between paired primary and metastatic CRCs. This result could be interpreted that the clones with the *TP53* mutation might be expanded through sustained tumor growth and metastasis or an additional genetic hit, resulting in loss of heterozygosity.

*SMAD4*, a downstream regulator in the TGF- $\beta$  signaling pathway in CRC, has been highlighted. In particular, inactivation of *SMAD4* has been associated with late stage or metastatic CRCs [18]. Recent work has also highlighted that *SMAD4* downregulation may occur in up to 60% of patients with metastatic CRC, which is significantly higher than the incidence of *SMAD4* mutations [19]. However, there was no statistical difference in the frequency of *SMAD4* mutations between primary and metastatic CRCs in our study.

*FBXW7* is a tumor suppressor gene and the frequency of mutation in CRCs has been reported at 6–10% [20–22]. It was recently reported that *FBXW7* mutations had significantly worse survival in metastatic CRCs [23]. In our study, recurrence occurred significantly early in patients with the *FBXW7* mutant than in patients with wild type, although the overall survival (OS) was not different according to *FBXW7* mutational status.

There was a lower prevalence of *BRAF* genes in the both primary (2%) and metastatic (3%) CRCs of the present study compared with TCGA (9%) and MSKCC (11%) data. According to a previous review, the prevalence of *BRAF*-mutated CRC is lower in Eastern Asian countries (0.7–11.4%) than in Western countries (3.7–20.6%) [24].

Only 4.9% of metastatic CRCs displayed an MSI-H genotype/phenotype, a frequency that is lower than that reported for primary CRCs but similar to the MSK cohort (4%) [3]. These findings could be explained by the lower tendency of MSI-H genotype/phenotype tumors to metastasize [25].

The limitation of this study is that we could not analyze the difference of frequency and concordant rate of genes between primary and metastatic CRCs according to MSI-status due to the too small size of the MSI-H genotype/phenotype tumors. Therefore, future studies, with a large cohort of MSI-H of metastatic CRCs, are needed to further investigate the difference of frequency and concordance in MSS and MSI-H group. A second limitation is that our study only examined the changes in DNA. The areas that are not well explained about the association with clinicopathological factors may be explained by other aspects of genetic variation, such as methylation and microenvironment.

#### 4. Materials and Methods

##### 4.1. Patients

A total of 142 patients with CRC and distant metastasis were enrolled in this study. Ninety-five cases were available for both primary and metastatic tumor tissues. Additional 47 metastatic only samples were included for mutational profiling of metastatic tumors and 4 primary only samples were used as a control for metastatic tumors. H&E stained slides were reviewed by pathologists and samples with tumor cellularity  $\geq 50\%$  were included to detect confident variants. Clinicopathological information, including age, sex, smoking history, and stage of cancer, was obtained retrospectively by reviewing medical records. Synchronous metastasis was defined as a metastatic disease at the time or within 6 months of the original diagnosis of CRC. Metachronous metastasis was defined as the absence of metastatic disease at the time of the initial diagnosis with metastatic disease developing later than 6 months from the original diagnosis. The study protocol was approved by the institutional review board of Konkuk University Medical Center (KUH1210052), and written informed consents were obtained from all patients.

##### 4.2. Targeted Sequencing & Identification of Clinically Significant Mutations

Custom panel including 170 cancer genes (Table S2) was used for mutational profiling. Sample preparation, sequencing procedures, and variant calling pipeline were almost same as the one described in our previous publication [26]. After variant calling, variant annotations were done using VEP [27] to annotate genes, variant type, gnomAD [28] minor allele frequency (MAF), and ClinVar [29] significance. Annotation with OncoKB [30] variant annotator (<https://github.com/oncokb/oncokb-annotator>, accessed on 1 September 2020) was also performed to check oncogenicities of variants.

To identify confident and clinically significant variants, we applied filtering criteria as followings: (1) protein altering variants including missense and nonsense SNVs, splice site mutations and indels, (2) rare variants with MAF of gnomAD total and east Asian  $\leq 0.1\%$ , (3) probable pathogenic mutations with ClinVar significance in one of 'Pathogenic', 'Likely pathogenic', or 'Drug response', or oncoKB annotation with one of 'Predicted oncogenic', 'Likely oncogenic', or 'Oncogenic'.

##### 4.3. Comparing with Public Data

To compare mutational frequencies of primary and metastatic tumors with public data, we downloaded files from cBioPortal. (<https://www.cbioportal.org/datasets>, accessed on 19 May 2021) Annotated variants of dataset 'Colorectal Adenocarcinoma (TCGA, Firehose Legacy)' ( $n = 223$ ) were compared with primary tumor group and 'Metastatic Colorectal Cancer (MSKCC, Cancer Cell 2018)' ( $n = 1134$ ) were compared with metastatic tumor group. To compare public data with our results under the same conditions, ClinVar and OncoKB annotation were performed for the two datasets and filtering steps by the criteria described above were applied.

##### 4.4. Molecular Findings for Microsatellite Instability (MSI)

We performed an MSI analysis on paraffin-embedded tissues to evaluate MSI status. The MSI status of the tumor samples was determined by using the five-marker Bethesda panel (BAT25, BAT26, D5S346, D2S123 and D17S250) [31]. Polymerase chain reaction (PCR)

products were run on a Qsep 100 DNA fragment analyzer (Bioptic Inc., Taiwan, China) and analyzed using Qsep 100 viewer (Bioptic Inc., Taiwan, China). Microsatellite instability was defined by the presence of different sized alleles in tumor DNA compared with the matched normal DNA sample. We classified the results into microsatellite instability-high (MSI-H), microsatellite instability-low (MSI-L) and microsatellite stable (MSS) in tumors according to Bethesda guidelines [32].

#### 4.5. Statistical Analysis

The association between mutational status and clinicopathological features was analyzed with the  $\chi^2$  test or Fisher's exact test. Paired two sample t-test was done by comparing the VAFs of 95 pairs of primary and metastatic tumors. Two sample proportion test was used in comparing mutational frequencies between groups. The overall survival (OS) was the primary endpoint for this study and was calculated from the date of surgery until the date of death. The Kaplan–Meier method was used to estimate the OS. *p*-value of less than 0.05 was considered to indicate a statistically significant difference. All analyses were carried out using Rex software version 3.0.3 (RexSoft Inc., Seoul, Korea).

## 5. Conclusions

In conclusion, our data is significant in comparing with primary and metastatic CRCs in pairs, as well as comparing two large public data, TCGA and MSKCC, representative of primary and metastatic CRCs, under the same conditions. To the author's knowledge, this is the first study to compare the genetic profiles of our cohort with that of the metastatic CRCs from MSKCC dataset. Comparative sequencing analysis between primary and metastatic CRCs revealed a high degree of genetic concordance in the main driver genes, especially, the current clinically actionable genes. Therefore, the genetic investigation of archived primary tumor samples with the challenges of obtaining adequate samples from metastatic sites appears to be sufficient for the application of cancer precision medicine in the metastatic setting.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijms22115561/s1>, Table S1: Full list of variants; Table S2: Gene list included in Custom Cancer Panel (170 cancer-related genes).

**Author Contributions:** S.E.L. designed the experiments, conducted the main experiments, and prepared the original draft; H.Y.P. and H.S.H. supervised and revised the manuscript; S.E.L. and H.Y.P. analyzed the data; H.S.H. and D.-Y.H. contributed clarifications and guidance on the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study protocol was approved by the institutional review board of Konkuk University Medical Center (KUH1210052).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This work was supported by Konkuk University Medical Center Research Grant 2018.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Vogelstein, B.; Fearon, E.R.; Hamilton, S.R.; Kern, S.E.; Preisinger, A.C.; Leppert, M.; Nakamura, Y.; White, R.; Smits, A.M.; Bos, J.L. Genetic alterations during colorectal-tumor development. *N. Engl. J. Med.* **1988**, *319*, 525–532. [[CrossRef](#)]
2. Cancer Genome Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **2012**, *487*, 330–337. [[CrossRef](#)]
3. Yaeger, R.; Chatila, W.K.; Lipsyc, M.D.; Hechtman, J.F.; Cercek, A.; Sanchez-Vega, F.; Jayakumaran, G.; Middha, S.; Zehir, A.; Donoghue, M.T.A.; et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell* **2018**, *33*, 125–136 e3. [[CrossRef](#)]

4. Jones, S.; Chen, W.D.; Parmigiani, G.; Diehl, F.; Beerenwinkel, N.; Antal, T.; Traulsen, A.; Nowak, M.A.; Siegel, C.; Velculescu, V.E.; et al. Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4283–4288. [[CrossRef](#)]
5. Vakiani, E.; Janakiraman, M.; Shen, R.; Sinha, R.; Zeng, Z.; Shia, J.; Cercek, A.; Kemeny, N.; D'Angelica, M.; Viale, A.; et al. Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J. Clin. Oncol.* **2012**, *30*, 2956–2962. [[CrossRef](#)] [[PubMed](#)]
6. Brannon, A.R.; Vakiani, E.; Sylvester, B.E.; Scott, S.N.; McDermott, G.; Shah, R.H.; Kania, K.; Viale, A.; Oschwald, D.M.; Vacic, V.; et al. Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. *Genome Biol.* **2014**, *15*, 454. [[CrossRef](#)] [[PubMed](#)]
7. Lim, B.; Mun, J.; Kim, J.H.; Kim, C.W.; Roh, S.A.; Cho, D.H.; Kim, Y.S.; Kim, S.Y.; Kim, J.C. Genome-wide mutation profiles of colorectal tumors and associated liver metastases at the exome and transcriptome levels. *Oncotarget* **2015**, *6*, 22179–22190. [[CrossRef](#)]
8. Williams, M.J.; Werner, B.; Barnes, C.P.; Graham, T.A.; Sottoriva, A. Identification of neutral tumor evolution across cancer types. *Nat. Genet.* **2016**, *48*, 238–244. [[CrossRef](#)]
9. Saito, T.; Niida, A.; Uchi, R.; Hirata, H.; Komatsu, H.; Sakimura, S.; Hayashi, S.; Nambara, S.; Kuroda, Y.; Ito, S.; et al. A temporal shift of the evolutionary principle shaping intratumor heterogeneity in colorectal cancer. *Nat. Commun.* **2018**, *9*, 2884. [[CrossRef](#)]
10. Sondka, Z.; Bamford, S.; Cole, C.G.; Ward, S.A.; Dunham, I.; Forbes, S.A. The COSMIC Cancer Gene Census: Describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* **2018**, *18*, 696–705. [[CrossRef](#)]
11. Choi, J.Y.; Choi, S.; Lee, M.; Park, Y.S.; Sung, J.S.; Chang, W.J.; Kim, J.W.; Choi, Y.J.; Kim, J.; Kim, D.S.; et al. Clinical Implication of Concordant or Discordant Genomic Profiling between Primary and Matched Metastatic Tissues in Patients with Colorectal Cancer. *Cancer Res. Treat.* **2020**, *52*, 764–778. [[CrossRef](#)] [[PubMed](#)]
12. Zehir, A.; Benayed, R.; Shah, R.H.; Syed, A.; Middha, S.; Kim, H.R.; Srinivasan, P.; Gao, J.; Chakravarty, D.; Devlin, S.M.; et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **2017**, *23*, 703–713. [[CrossRef](#)] [[PubMed](#)]
13. Priestley, P.; Baber, J.; Lolkema, M.P.; Steeghs, N.; de Bruijn, E.; Shale, C.; Duyvesteyn, K.; Haidari, S.; van Hoeck, A.; Onstenk, W.; et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **2019**, *575*, 210–216. [[CrossRef](#)]
14. Vignot, S.; Lefebvre, C.; Frampton, G.M.; Meurice, G.; Yelensky, R.; Palmer, G.; Capron, F.; Lazar, V.; Hannoun, L.; Miller, V.A.; et al. Comparative analysis of primary tumour and matched metastases in colorectal cancer patients: Evaluation of concordance between genomic and transcriptional profiles. *Eur. J. Cancer* **2015**, *51*, 791–799. [[CrossRef](#)]
15. Cross, W.; Graham, T.A.; Wright, N.A. New paradigms in clonal evolution: Punctuated equilibrium in cancer. *J. Pathol.* **2016**, *240*, 126–136. [[CrossRef](#)]
16. Bhullar, D.S.; Barriuso, J.; Mullamitha, S.; Saunders, M.P.; O'Dwyer, S.T.; Aziz, O. Biomarker concordance between primary colorectal cancer and its metastases. *EBioMedicine* **2019**, *40*, 363–374. [[CrossRef](#)]
17. Xie, Y.H.; Chen, Y.X.; Fang, J.Y. Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct. Target. Ther.* **2020**, *5*, 22. [[CrossRef](#)]
18. Miyaki, M.; Iijima, T.; Konishi, M.; Sakai, K.; Ishii, A.; Yasuno, M.; Hishima, T.; Koike, M.; Shitara, N.; Iwama, T.; et al. Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene* **1999**, *18*, 3098–3103. [[CrossRef](#)]
19. Losi, L.; Luppi, G.; Benhattar, J. Assessment of K-ras, Smad4 and p53 gene alterations in colorectal metastases and their role in the metastatic process. *Oncol. Rep.* **2004**, *12*, 1221–1225. [[CrossRef](#)]
20. Kemp, Z.; Rowan, A.; Chambers, W.; Wortham, N.; Halford, S.; Sieber, O.; Mortensen, N.; von Herbay, A.; Gunther, T.; Ilyas, M.; et al. CDC4 mutations occur in a subset of colorectal cancers but are not predicted to cause loss of function and are not associated with chromosomal instability. *Cancer Res.* **2005**, *65*, 11361–11366. [[CrossRef](#)]
21. Akhoondi, S.; Sun, D.; von der Lehr, N.; Apostolidou, S.; Klotz, K.; Maljukova, A.; Cepeda, D.; Fiegl, H.; Dafou, D.; Marth, C.; et al. FBXW7/hCDC4 is a general tumor suppressor in human cancer. *Cancer Res.* **2007**, *67*, 9006–9012. [[CrossRef](#)]
22. Malapelle, U.; Pisapia, P.; Sgariglia, R.; Vigliar, E.; Biglietto, M.; Carlomagno, C.; Giuffrè, G.; Bellecicine, C.; Troncone, G. Less frequently mutated genes in colorectal cancer: Evidences from next-generation sequencing of 653 routine cases. *J. Clin. Pathol.* **2016**, *69*, 767–771. [[CrossRef](#)]
23. Korphaisarn, K.; Morris, V.K.; Overman, M.J.; Fogelman, D.R.; Kee, B.K.; Raghav, K.P.S.; Manuel, S.; Shureiqi, I.; Wolff, R.A.; Eng, C.; et al. FBXW7 missense mutation: A novel negative prognostic factor in metastatic colorectal adenocarcinoma. *Oncotarget* **2017**, *8*, 39268–39279. [[CrossRef](#)]
24. Kim, J.H.; Kang, G.H. Molecular and prognostic heterogeneity of microsatellite-unstable colorectal cancer. *World J. Gastroenterol.* **2014**, *20*, 4230–4243. [[CrossRef](#)]
25. Koopman, M.; Kortman, G.A.; Mekenkamp, L.; Ligtenberg, M.J.; Hoogerbrugge, N.; Antonini, N.F.; Punt, C.J.; van Krieken, J.H. Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br. J. Cancer* **2009**, *100*, 266–273. [[CrossRef](#)]
26. Lee, S.E.; Lee, H.S.; Kim, K.Y.; Park, J.H.; Roh, H.; Park, H.Y.; Kim, W.S. High prevalence of the MLH1 V384D germline mutation in patients with HER2-positive luminal B breast cancer. *Sci. Rep.* **2019**, *9*, 10966. [[CrossRef](#)]
27. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biol.* **2016**, *17*, 122. [[CrossRef](#)]

28. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [[CrossRef](#)]
29. Landrum, M.J.; Lee, J.M.; Benson, M.; Brown, G.R.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **2018**, *46*, D1062–D1067. [[CrossRef](#)]
30. Chakravarty, D.; Gao, J.; Phillips, S.M.; Kundra, R.; Zhang, H.; Wang, J.; Rudolph, J.E.; Yaeger, R.; Soumerai, T.; Nissan, M.H.; et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis. Oncol.* **2017**. [[CrossRef](#)]
31. Boland, C.R.; Thibodeau, S.N.; Hamilton, S.R.; Sidransky, D.; Eshleman, J.R.; Burt, R.W.; Meltzer, S.J.; Rodriguez-Bigas, M.A.; Fodde, R.; Ranzani, G.N.; 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.* **1998**, *58*, 5248–5257. [[PubMed](#)]
32. Umar, A.; Boland, C.R.; Terdiman, J.P.; Syngal, S.; de la Chapelle, A.; Ruschoff, J.; Fishel, R.; Lindor, N.M.; Burgart, L.J.; Hamelin, R.; et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J. Natl. Cancer Inst.* **2004**, *96*, 261–268. [[CrossRef](#)] [[PubMed](#)]