

RESEARCH ARTICLE

A Pathway-Centric Survey of Somatic Mutations in Chinese Patients with Colorectal Carcinomas

Chao Ling^{1,2,3‡}, Lin Wang^{4‡}, Zheng Wang⁴, Luming Xu⁴, Lifang Sun⁴, Hui Yang¹, Wei-Dong Li³, Kai Wang^{1,5,6*}

1 Zilkha Neurogenetic Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, 90089, United States of America, **2** Laboratory of Clinical Genetics, Peking Union Medical College Hospital, Beijing, 100730, China, **3** Research Center of Basic Medical Sciences, Tianjin Medical University, Tianjin, 300070, China, **4** Center for Tissue Engineering and Regenerative Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, 430022, China, **5** USC/Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, 90089, United States of America, **6** Department of Psychiatry, Keck School of Medicine, University of Southern California, Los Angeles, CA, 90089, United States of America

‡ These authors contributed equally to this work.

* kaiwang@usc.edu



OPEN ACCESS

Citation: Ling C, Wang L, Wang Z, Xu L, Sun L, Yang H, et al. (2015) A Pathway-Centric Survey of Somatic Mutations in Chinese Patients with Colorectal Carcinomas. PLoS ONE 10(1): e0116753. doi:10.1371/journal.pone.0116753

Academic Editor: Hiromu Suzuki, Sapporo Medical University, JAPAN

Received: August 1, 2014

Accepted: December 11, 2014

Published: January 24, 2015

Copyright: © 2015 Ling et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Raw sequence data was submitted to sequence read archive (SRA). The accession number is SRP045337.

Funding: This project was supported by start-up funds from the Zilkha Neurogenetic Institute at the University of Southern California. KW received the funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors confirm that Dr. Kai Wang is one of PLOS ONE Editorial Board

Abstract

Previous genetic studies on colorectal carcinomas (CRC) have identified multiple somatic mutations in four candidate pathways (*TGF-β*, *Wnt*, *P53* and *RTK-RAS* pathways) on populations of European ancestry. However, it is under-studied whether other populations harbor different sets of hot-spot somatic mutations in these pathways and other oncogenes. In this study, to evaluate the mutational spectrum of novel somatic mutations, we assessed 41 pairs of tumor-stroma tissues from Chinese patients with CRC, including 29 colon carcinomas and 12 rectal carcinomas. We designed Illumina Custom Amplicon panel to target 43 genes, including genes in the four candidate pathways, as well as several known oncogenes for other cancers. Candidate mutations were validated by Sanger sequencing, and we further used SIFT and PolyPhen-2 to assess potentially functional mutations. We discovered 3 new somatic mutations in gene *APC*, *TCF7L2*, and *PIK3CA* that had never been reported in the COSMIC or NCI-60 databases. Additionally, we confirmed 6 known somatic mutations in gene *SMAD4*, *APC*, *FBXW7*, *BRAF* and *PTEN* in Chinese CRC patients. While most were previously reported in CRC, one mutation in *PTEN* was reported only in malignant endometrium cancer. Our study confirmed the existence of known somatic mutations in the four candidate pathways for CRC in Chinese patients. We also discovered a number of novel somatic mutations in these pathways, which may have implications for the pathogenesis of CRC.

members, but this does not alter the authors' adherence to PLOS ONE Editorial policies and criteria.

Introduction

Colorectal carcinoma (CRC) is the third most commonly diagnosed cancer in males yet the second in females, and the World Health Organization GLOBOCAN database reported approximately 1.2 million new cancer cases and more than 608,000 deaths globally in 2008 [1]. The molecular genetic mechanism underlying CRC tumorigenesis appear to be multifactorial [2, 3]. Substantial studies discovered that CRC arises as a result of the accumulation of genetic and epigenetic alterations [4], including loss of genomic stability, microsatellite instability, aberrant DNA methylation and DNA repair defects, which highlight the heterogeneity of the disease, besides the controllable contributors such as dietary and lifestyle factors [5, 6].

Earlier generation of genome-wide association studies (GWAS) identified 8q24, 11q23, 18q21, 10p14 and 8q23.3 as susceptibility loci of CRC [7–9], but more recently, over 20 additional loci such as rs647161 (5q31.1), rs2423279 (20p12.3) and rs10774214 (12p13.32) were discovered to be associated with CRC [10]. These results from GWAS confirmed the heritability of CRC initiation [11, 12]. On the other hand, rare germline mutation of the *APC* gene was found in both sporadic CRC [13] and familial adenomatous polyposis [14], and other studies implied that high-penetrant germline mutations of the *MSH2* and *MLH1* gene [15] may account for hereditary nonpolyposis colorectal cancer [16].

In addition to germline variants that confer susceptibility to CRCs, genome-wide analysis of mutations in CRCs has identified acquired somatic mutations in several hundred genes, with an average of 80 mutations in any single patient with CRC [2, 17]. Importantly, somatic mutations in *TGF- β* , *Wnt*, *APC*, *TP53* and *RTK-RAS* pathways were identified to be important contributors to CRC development [18]. Among them, *APC* was proposed to be a gatekeeper to regulate epithelial cells to develop into the adenoma and even carcinoma cells [19], yet alteration of *TP53* was proposed to be a relatively late event in the development of colorectal tumor. Without *P53*-mediation, cell apoptosis signal loses exertion of cell cycle control and may result in the progression from adenoma to malignant tumor [20]. Recently, with the development of next generation sequencing techniques, a series of somatic point mutations and rearrangements of oncogenes or tumor suppressor genes (TSG) underlying the tumorigenesis of CRC were confirmed, prominently the genes *APC* [21], *K-ras* [22], and *TP53* [23]. Additionally, somatic mutations in genes *SMAD2*, *SMAD3* and *SMAD4* had also been identified in previous CRC studies [24].

Although the genetic alterations associated with CRC had been extensively studied, the incidence of CRC varies widely between worldwide populations. Statistically, around 60% of cases are diagnosed in the developed world [25], and the highest incidence rates are found in North America and Europe, Australia and New Zealand, while the lowest rates are in Africa and South-Central Asia [1]. Besides environmental factors, this implies that different genetic background may play a role in CRC susceptibility, and that different patterns of somatic mutations may contribute to CRC progression. Additionally, a recent study identified novel DNA variants of *hMLH1* and *hMSH2* genes in Chinese CRC patients [26]. Therefore, in this study, we selected 41 paired tumor-stroma tissues from the Chinese CRC patients representing all grades, to explore the patterns of somatic mutations in the four candidate pathways and in oncogenes for other cancer types, and to complete the understanding of somatic mutation in CRC patients.

Materials and Methods

Sample collection

This study was conducted according to the Helsinki human subject doctrine and was approved by the Wuhan Union Hospital review board and Ethics Committee; informed consent was

signed and obtained from all participants for tissue specimen collection and subsequent analysis.

We collected 41 pairs of colorectal carcinomas (CRC) and their surrounding normal stroma tissue samples (Table 1) from Wuhan Union Hospital, Wuhan, China. All CRCs were diagnosed based on pathological examination and laboratory evaluation and classified with TNM staging system. Genomic DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) samples. Five 6- μm -thick serial sections were cut from each paraffin block and collected in Eppendorf tubes. Paraffin was removed by xylene, DNA was extracted with QIAamp DNA FFPE kit (Qiagen, CA, USA) according to the manufacturer's recommendations. The DNA was quantified by NanoDrop ND-1000 spectrophotometer and analyzed with agarose electrophoresis. The 260/280 ratios range from 1.7~2.0, yet gel electrophoresis indicated generally acceptable quality for these samples.

Design of custom gene capture panel

We used the Illumina TruSeq Custom Amplicon assay, which is one of the fastest and easiest-to-use multiplexed amplicon assay optimized for the MiSeq system, to capture and enrich exons from candidate genes. The assay allows researchers to sequence up to 1,536 amplicons in a single reaction using a simple workflow, and it is known to work well with FFPE-derived DNA samples. We selected 43 candidate genes, including all genes previously reported in the OncoMap project [27, 28], as well as candidate genes in the *TGF- β* , *Wnt*, *P53* and *RTK-RAS* pathway, respectively. Detailed description of these genes was given in Table 2 and S5 Table. In total, 1,484 custom amplicons of 43 genes were designed from these genes and were synthesized by Illumina as Custom Amplicon assay.

Sequencing library preparation

For DNA library preparation, ~300 ng whole genome DNA was used as input. First, a custom pool containing upstream and downstream oligos specific to the targeted regions of interest (S5 Table) was hybridized to the genomic DNA samples. Second, a DNA polymerase extended from the upstream oligo through the target region, followed by ligation to the 5' end the downstream oligo using a DNA ligase. Then, the extension-ligation products were amplified using primers that add index sequences for sample multiplexing (i5 and i7) as well as common adapters required for cluster generation (P5 and P7). Next, PCR was performed on the thermal cycler using the program of 95°C for 3 minutes, and 35 cycles of 95°C for 30 seconds, 66°C for 30 seconds, 72°C for 60 seconds and 72°C for 5 minutes. After the PCR amplification, AMPure XP beads were used to purify the PCR products, and then TruSeq normalization beads were used to normalize the quantity of each library to ensure more equal library representation in the pooled sample. DNA library was prepared using Illumina TruSeq Amplicon panel; 5 μl of each library was pooled and Kapa q-PCR was used for concentration checking before MiSeq sequencing.

Statistical analysis and somatic mutation detection

We used the Illumina Experiment Manager to define the sample sheet, and the sample library was loaded into the MiSeq reagent cartridge for automated cluster generation and sequencing. The library quality control (QC) report was generated on the instrument using MiSeq Reporter software. The Illumina somatic variant analysis program align the reads to human reference genome hg19 with banded Smith-Waterman algorithm, and then we used VarScan2 [29] for calling somatic variants by the paired tumor-stroma mode with default program parameters: min coverage: 8x for normal, 6x for tumor, somatic P-value was 0.05. We selected $P < 5 \times 10^{-5}$ as the

Table 1. Patients with CRCs enrolled in this study.

Patient	Gender	Age	Classification	Location	TNM
5182	M	23Y	Familial adenomatous polyposis and cancer	whole colon	Tis
5862	M	81Y	Well differentiated adenocarcinoma	rectum	T1
6864	M	61Y	Moderately differentiated adenocarcinoma	ascending colon	T1
8026	F	79Y	Moderately differentiated adenocarcinoma	descending colon	T1
6679	F	44Y	Moderately differentiated adenocarcinoma	rectum	T2
6840	F	66Y	Moderately differentiated adenocarcinoma	rectum	T2
7491	F	65Y	Poorly differentiated adenocarcinoma	sigmoid colon	T2
3534	F	64Y	Moderately differentiated adenocarcinoma	ascending colon	T3
3636	M	47Y	Well differentiated adenocarcinoma	sigmoid colon	T3
4950	M	59Y	Moderately differentiated adenocarcinoma	sigmoid colon	T3
5181	M	78Y	Moderately differentiated adenocarcinoma	transverse colon	T3
5457	M	45Y	Moderately differentiated adenocarcinoma	descending colon	T3
5550	F	71Y	Well differentiated adenocarcinoma	rectum	T3
5557	F	58Y	Poorly differentiated adenocarcinoma	transverse colon	T3
6472	F	48Y	Moderately differentiated adenocarcinoma	sigmoid colon	T3
6515	M	59Y	Moderately differentiated adenocarcinoma	rectum	T3
6856	M	61Y	Moderately differentiated adenocarcinoma	sigmoid colon	T3
6935	M	77Y	Poorly differentiated adenocarcinoma	ascending colon	T3
6990	F	81Y	Moderately differentiated adenocarcinoma	descending colon	T3
6992	M	52Y	Poorly differentiated adenocarcinoma	ascending colon	T3
7411	M	66Y	Poorly differentiated adenocarcinoma	sigmoid colon	T3
7507	M	63Y	Moderately differentiated adenocarcinoma	sigmoid colon	T3
7573	M	65Y	Poorly differentiated adenocarcinoma	rectum	T3
8096	M	84Y	Moderately differentiated adenocarcinoma	sigmoid colon	T3
8103	F	51Y	Mucinous carcinoma	colon	T3
9101	M	78Y	Moderately differentiated adenocarcinoma	colon	T3
4996	M	79Y	Moderately differentiated adenocarcinoma	ileocecal area	T4a
5569	M	51Y	Moderately differentiated adenocarcinoma	ascending colon	T4a
6020	M	49Y	Poorly differentiated adenocarcinoma	rectum	T4a
6194	M	46Y	Mucinous carcinoma	rectum	T4a
6266	M	89Y	Moderately differentiated adenocarcinoma	sigmoid colon	T4a
6284	M	66Y	Poorly differentiated adenocarcinoma	rectum	T4a
6566	F	64Y	Moderately differentiated adenocarcinoma	rectum	T4a
7506	M	23Y	Moderately differentiated adenocarcinoma	descending colon	T4a
7510	M	78Y	Moderately differentiated adenocarcinoma	rectum	T4a
7753	M	72Y	Moderately differentiated adenocarcinoma	sigmoid colon	T4a
8280	M	60Y	Moderately differentiated adenocarcinoma	transverse colon	T4a
8993	M	54Y	Moderately differentiated adenocarcinoma	colon	T4a
5470	M	59Y	Moderately differentiated adenocarcinoma	sigmoid colon	T4b
5755	F	66Y	Moderately differentiated adenocarcinoma	rectum	T4b
6591	M	51Y	Mucinous carcinoma	transverse colon	T4b

Gender: M, male; F, female. Y: Year.

TNM: Tis, carcinoma in situ; T1, tumor invades submucosa; T2, tumor invades muscularispropria; T3, tumor invades through the muscularispropria into the pericolorectal tissues; T4a, tumor penetrates to the surface of the visceral peritoneum; T4b, tumor directly invades or is adherent to other organs or structures.

doi:10.1371/journal.pone.0116753.t001

Table 2. Genes examined in this study.

Gene	Official full name	Gene	Official full name
<i>ABL1</i>	c-abl oncogene 1	<i>KRAS</i>	Kirsten rat sarcoma viral oncogene homolog
<i>AKT1</i>	v-akt murine thymoma viral oncogene homolog 1	<i>MET</i>	met proto-oncogene
<i>AKT2</i>	v-akt murine thymoma viral oncogene homolog 2	<i>MLH1</i>	mutL homolog 1
<i>APC</i>	adenomatous polyposis coli	<i>MYC</i>	v-myc avian myelocytomatosis viral oncogene homolog
<i>ARID1A</i>	AT rich interactive domain 1A (SWI-like)	<i>NF1</i>	neurofibromin 1
<i>BRAF</i>	v-raf murine sarcoma viral oncogene homolog B	<i>NF2</i>	neurofibromin 2 (merlin)
<i>CDK4</i>	cyclin-dependent kinase 4	<i>NRAS</i>	neuroblastoma RAS viral (v-ras) oncogene homolog
<i>CDKN2A</i>	cyclin-dependent kinase inhibitor 2A	<i>PDGFRA</i>	platelet-derived growth factor receptor, alpha polypeptide
<i>CSF1R</i>	colony stimulating factor 1 receptor	<i>PIK3CA</i>	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha
<i>CTNNB1</i>	catenin (cadherin-associated protein), beta 1, 88kDa	<i>PTEN</i>	phosphatase and tensin homolog
<i>EGFR</i>	epidermal growth factor receptor	<i>RB1</i>	retinoblastoma 1
<i>ERBB2</i>	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2	<i>RET</i>	ret proto-oncogene
<i>FAM123B</i>	family with sequence similarity 123B	<i>SMAD2</i>	SMAD family member 2
<i>FBXW7</i>	F-box and WD repeat domain containing 7, E3 ubiquitin protein ligase	<i>SMAD3</i>	SMAD family member 3
<i>FGFR1</i>	fibroblast growth factor receptor 1	<i>SMAD4</i>	SMAD family member 4
<i>FGFR2</i>	fibroblast growth factor receptor 2	<i>SOX9</i>	SRY (sex determining region Y)-box 9
<i>FGFR3</i>	fibroblast growth factor receptor 3	<i>SRC</i>	v-src avian sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog
<i>FLT3</i>	fms-related tyrosine kinase 3	<i>STK11</i>	serine/threonine kinase 11
<i>HRAS</i>	Harvey rat sarcoma viral oncogene homolog	<i>TCF7L2</i>	transcription factor 7-like 2 (T-cell specific, HMG-box)
<i>JAK2</i>	Janus kinase 2	<i>TP53</i>	tumor protein p53
<i>JAK3</i>	Janus kinase 3	<i>VHL</i>	von Hippel-Lindau tumor suppressor, E3 ubiquitin protein ligase
<i>KIT</i>	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog		

doi:10.1371/journal.pone.0116753.t002

threshold for highly reliable somatic mutation detection. The WANNONAR web server [30, 31] was used for variants annotation, and those variants observed in the 1000 Genome Project, dbSNP138, or ESP6500 data sets were excluded since they are likely to be germline variants. Additionally, synonymous variants were removed since they are unlikely to change protein function. Detailed data filtering procedure is described in Fig. 1. Raw sequence data was submitted to sequence read archive (SRA), and the accession number is SRP045337.

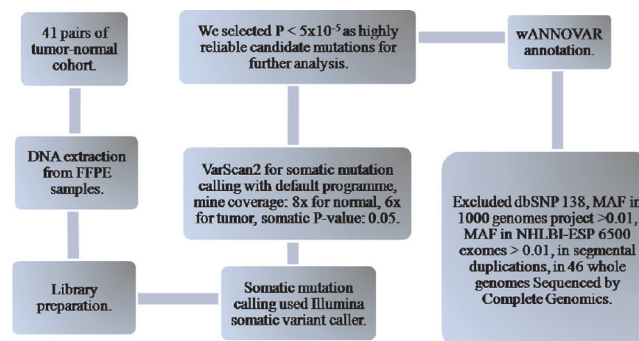


Figure 1. Flowchart of the data analysis strategy.

doi:10.1371/journal.pone.0116753.g001

Sanger validation of selected mutations

For selected mutations prioritized for Sanger validation, we designed PCR primers with Primer-BLAST [32] (S6 Table), and validated the presence of candidate mutations by Sanger sequencing on both tumor and stroma samples [33]. The sequencing traces from tumor and stroma tissues were manually visualized in ABI Sequence Scanner software to confirm the mutation calls.

Results

Identification of somatic mutations

Using the Illumina MiSeq platform, we sequenced 43 genes (Table 2) in 41 pairs of tumor-stroma samples from the Chinese patients with CRCs. The average age of the patients is 61.8 years old (standard deviation: 14.9), and the male:female ratio is 2.4:1. These samples have various TNM stages, ranging from Tis to T4b (Table 1). From raw sequence data, we assessed the data quality by FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), and calculated the total average coverage was 182x for subjects, with 95% regions having an average coverage over 20x (S1 Table). Given the relatively small number of reads, the alignment was performed by the Smith-Waterman algorithm [34] per Illumina software for MiSeq sequencer, which performs local sequence alignments to determine similar regions between two sequences. Alignments that include more than three indels were filtered from the results, and were not included in the alignment (BAM) files for variant calling. We used VarScan2 [29] paired-sample mode to explore somatic mutations with default somatic calling parameters, and mutations with P-value $< 5 \times 10^{-5}$ between the carcinoma-stroma comparison were regarded as high confidence variants. In total, we identified 212 candidate somatic mutations (S2 Table); 187 were reported as novel mutations (S3 Table), and 25 of them were previously reported in CRC or endometrium carcinoma (S4 Table). These candidate mutations occurred on 35 candidate genes from 28 CRC samples. Among all of the candidate mutations, we found that the gene *APC* was the most frequently mutated gene in our CRC patients, with 44 mutations in 13 samples. Among all the samples, one patient (ID7506) has the most number of mutations and he is indeed a young patient who is only 23 years old with the CRC stage of T4a. We selected the recurrent (≥ 2 samples) mutations with both low P value ($< 5 \times 10^{-5}$) and sufficiently high coverage (30x) for Sanger validation.

Somatic mutations in candidate pathways

Based on our analytical strategy (Fig. 1), we identified multiple genes in the *TGF β* , *Wnt* signaling, *P53* and *RTK-RAS* pathways that carried somatic mutations in a fraction of the cases (Fig. 2). Overall, 25 point mutations in 11 genes were reported in previous studies and were documented in the COSMIC v67 database [35] (S4 Table). The *APC* gene in *Wnt* pathway has relatively high levels of somatic mutations, compared to genes in the *TGF β* , *P53* and *RTK-RAS* pathways (S1 Table). We confirmed the recurrent somatic mutations in the *SMAD4* gene, including a recurrent point mutation of (c.1082G>A, p.R361H) [36] in 2 of our CRC cases. A nonsense point mutation in the tumor suppressor gene *APC* (c.4013C>G, p.S1338X) was previously reported in colon cancer [37], but was found in one of our cases with rectum cancer. Reported mutations in gene *FBXW7* (c.1514G>A, p.R505H) [38] and *BRAF* (c.1396G>A, p.G466R) (COSMIC) were also identified in our study. Meanwhile, a somatic mutation in the *PTEN* gene (c.415T>G, p.L139V) was previously found in endometrium cancer [39], but was identified in one of our cases with colon cancer. All these mutations have been validated by Sanger sequencing (Fig. 3A).

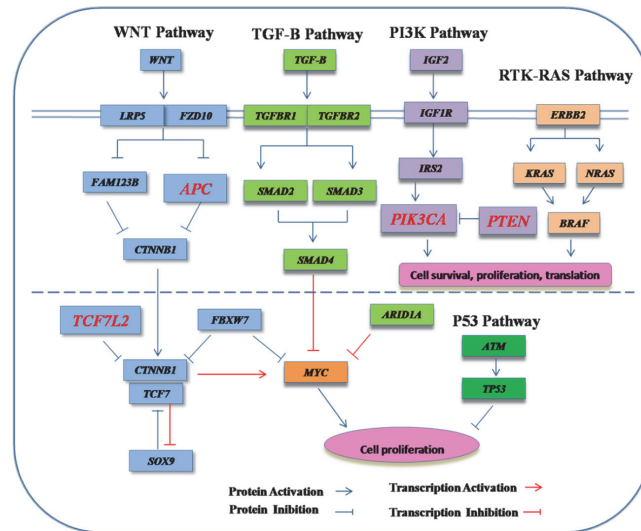


Figure 2. Diagram of the four major pathways evaluated in this study.

doi:10.1371/journal.pone.0116753.g002

Novel somatic mutations in CRC

Besides known somatic mutations identified from candidate pathways, we were also interested to assess whether novel somatic mutations can be identified from the Chinese CRC patients. In total, we validated 4 mutations in 3 genes that have not been reported in public databases before. One of them occurred in gene *APC* (c.1273G>T, p.E425X) which is a non-sense mutation, while one resides in the *TCF7L2* gene (c.2T>C, p.M1T) which changes the start codon and potentially results in non-functional protein (Fig. 3B). Both of the two genes were members of *Wnt* pathway. To our knowledge these mutations had never been detected in CRC or any other cancers. Additionally, we also found a novel mutation in gene *PIK3CA* (c.3001C>G, p.L1001V), which the PolyPhen-2 and SIFT software predicted as a deleterious mutation; interestingly, we found a previous study which reported a different mutation of c.3001C>A in thyroid carcinoma [40], which may suggest that this is a hot spot for mutations in different carcinomas.

Discussion

In the current study, we identified somatic mutations from genes in four critical candidate pathways and several known oncogenes in Chinese CRC patients. Multiple genes in the four important pathways of *TGFβ*, *Wnt*, *RAS* and *TP53* were mutated [38], and our results largely confirmed previous studies reporting that these pathways carry important driver mutations for initiation and progression of CRC. Additionally, we identified 4 mutations in 3 genes that have not been reported in public databases, and validated the presence of these mutations. The results of this study improved the understanding of somatic mutations in CRC, and may contribute to identify potential therapeutic targets.

We identified point mutations of the *SMAD4* gene [41] in 2 of stage T3 CRC cases, suggesting that this recurrent mutation may facilitate CRC progression. The *SMAD4* gene had been recognized as a candidate cancer gene, which plays a pivotal role in signal transduction of *TGFβ* pathway by mediating transcriptional activation of target genes [42]. Thus, the somatic mutation would affect the function of the *SMAD4* gene; however, the underlying mechanism why these mutations leads to alteration of gene function still needs additional studies. Gene

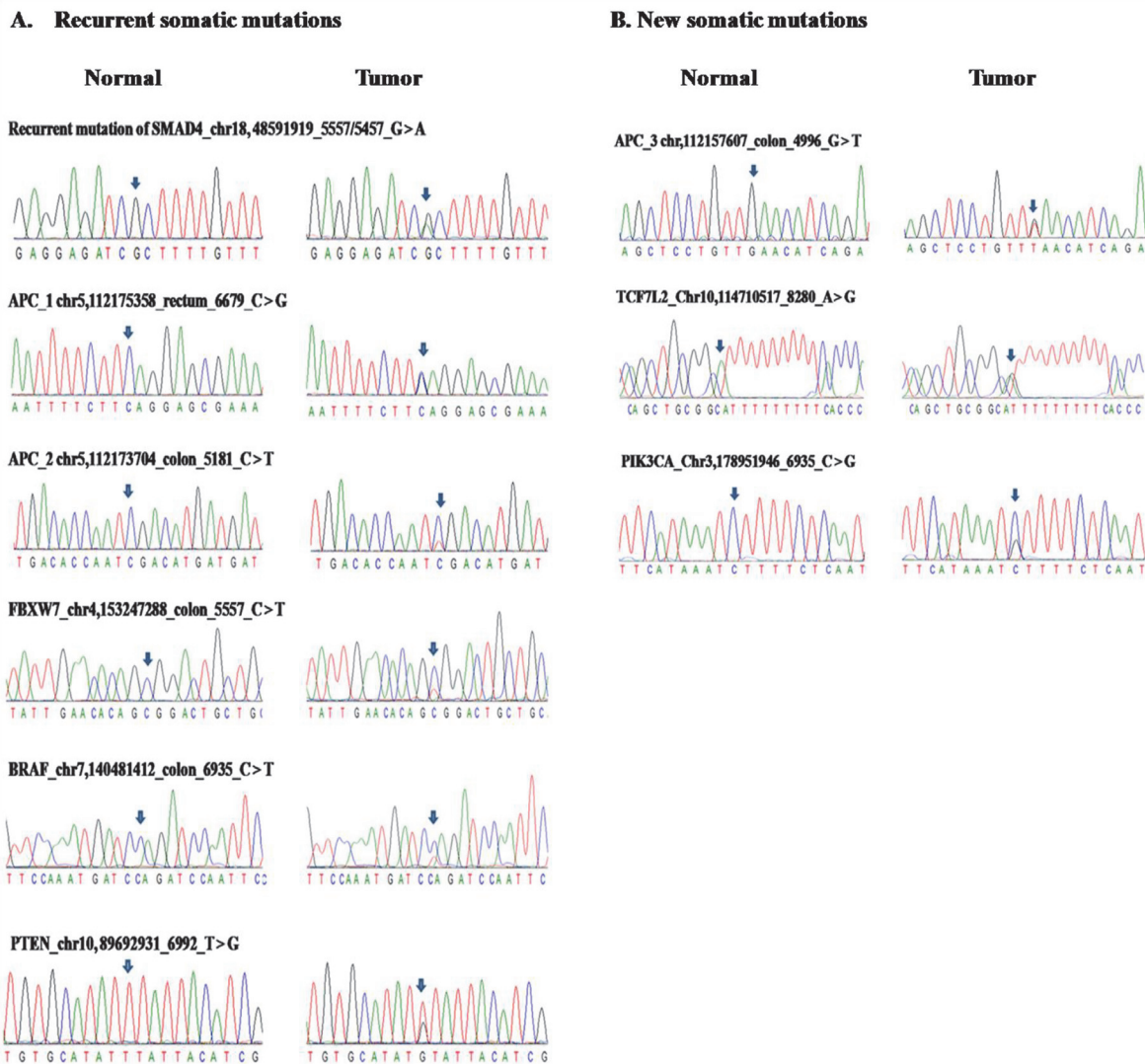


Figure 3. Sanger sequencing validated somatic mutations. A, somatic mutations which had been reported in CRC or other cancers. B, novel somatic mutations that discovered in our CRC patients.

doi:10.1371/journal.pone.0116753.g003

FBXW7 was involved in *TGFβ* and *Wnt* pathway, so mutation of gene *FBXW7* may largely account for the CRC tumorigenesis [38]. Additionally, the mutation of the gene *BRAF* in RTK-RAS pathway could be found in CRC, although it was frequently mutated in thyroid carcinoma [43]. This may imply that RTK-RAS pathway maybe common in contributions of tumorigenesis of CRC and thyroid carcinoma.

The *PTEN* gene can inactivate the *PI3K* pathway which induces cell survival and proliferation, and was previously found to have a high mutation rate in young (<50 years old) CRC patients [44]. Additional study reported that somatic genetic and epigenetic inactivation of *PTEN* was involved in as high as 93% of sporadic endometrial carcinomas, which was the most frequent extracolonic cancer in patients with hereditary nonpolyposis colon cancer (HNPCC) syndrome, suggesting that *PTEN* may play a significant pathogenic role in both HNPCC and sporadic endometrial carcinogenesis [45]. In our study, we identified *PTEN* nonsynonymous

somatic mutation in a 52 years old male patient with sporadic nonpolyposis colon cancer. Interestingly, the same point mutation occurred in our case was previously reported only in endometrial carcinoma [39], yet we confirmed that this mutation is a recurrent mutation in Chinese CRC patients. In the PI3K pathway, a novel mutation in the *PIK3CA* gene was validated in a 77 years old male patient. Previous studies reported that *PIK3CA* mutation frequently occurred in breast and hepatocellular carcinomas [46, 47]. In colorectal carcinoma, somatic *PIK3CA* mutation was recognized as a molecular biomarker that predicts response to aspirin therapy [48]. So our study may provide a novel candidate point mutation, which could assist in the prediction of aspirin therapy and avoid mistreatment for CRC.

Previous study suggested that somatic mutation of the *APC* gene occurred in both sporadic colorectal carcinoma and familial adenomatous polyposis [49, 50]. Mutations in the *APC* gene could be found in 81% of colorectal cancer cell lines [51], indicating the critical role of *APC* in CRC tumorigenesis. In this study, we included one FAP patient (sample 5181), because this patient had a colon cancer, which was developed from FAP, and here we are interested to know if there is any specific mutation in this single sample. Then we identified *APC* somatic mutation in this patient, and this imply that mutation of *APC* may contribute to tumorigenesis, but we cannot tell if it is an initiating event for FAP or colon cancer. The *APC* gene harbors the most number of somatic mutations in our study; although the frequency of *APC* mutations is around 30%, which is lower than previous studies on European populations, the results nevertheless confirmed that somatic mutation in *APC* is a key cancer driver for CRC. Due to small sample size, the differences are not significant compared to previous studies, and we combined patients from all stages together which may further dilute the signal for *APC* mutations. Interestingly, a recent exome-wide study on single colon cells illustrated that colon cancer could be of a bioclonal origin: One is a major tumor clone harboring *APC* and *TP53* mutation, and another one is a minor tumor clone without *APC* and *TP53* mutation [52]. Nevertheless, due to the limited sample size, the use of different technical platforms as previous studies and the paucity of similar studies in Chinese populations, we still could not confirm whether the mutation rate for *APC* differs between Chinese population and other ethnic groups.

We also identified previously unreported somatic mutations in gene *TCF7L2* in the Chinese CRC patients. Previous study suggested that *TCF7L2* may act as a feedback repressor of beta-catenin/*TCF4* genes, and may cooperate with *APC* to suppress malignant transformation of epithelial cells [53]. Here, the identification of novel somatic mutations in the *Wnt* pathway further suggested its role in tumor initiation and progression. Based on our analysis on pathways, we confirmed that mutations in different genes in the same pathway may lead to pathway dysfunction, which may result in cell proliferation and even metastasis.

Additionally, based on Table 2 and Fig. 2, apart from *APC*, genes that are frequently mutated include *ARID1A*, *PIK3CA*, *SMAD4* and *TCF7L2*. These genes were in *WNT*, *TGFβ* and *PI3K* pathways, which were reported to play an important role in CRC tumorigenesis and differentiation. Despite the small sample size, we further analyzed co-occurrence patterns, and found that patients with *APC* mutation may harbor mutations in these canonical pathways as well as outside of these canonical pathways (such as *ABL1*). For example, we found one poorly differentiated adenocarcinoma (Sample 6992) which co-harbored *APC* and *TP53* mutations (S2 Table). Additionally, we found that sample 6935 had co-occurring mutations in *PI3K* (*PIK3CA*) and *RTK-RAS* (*BRAF*) pathways, and the patient suffered a poorly differentiated adenocarcinoma in the ascending colon. Furthermore, sample 5557 had two somatic mutations in *Wnt* (*FBXW7*) and *TGFβ* (*SMAD4*) pathways, and this patient also suffered a poorly differentiated adenocarcinoma in transverse colon. The TNM grades in all of the 3 patients were T3, which implied that mutations in different genes may lead to dysfunction in multiple pathways, which may contribute to late-stage tumor progression and invasion.

Based on analysis on the data in [Table 1](#) and [Fig. 2](#), there was no significant association between somatic mutations and the anatomy location of CRC carcinoma, as somatic mutations could occur in any part of colon or rectum. Meanwhile, based on differences in TNM stage, we did not found significant association between somatic mutation of any gene with the stage of TNM. We want to stress that previous studies on pathway analysis mainly focused on populations of European ancestry [[54](#), [55](#)], with less emphasis on other ethnicity groups. Our study investigated the somatic mutation in Chinese CRC patients; however, we still could not confirm whether these novel mutations were specific in Chinese patients without further larger-scale study in a different ethnic population. Another explanation for the novel mutations might be the use of different technical platforms and analytical methods for somatic mutation detection: Illumina custom amplicon panel, MiSeq sequencer and VarScan2 pipeline.

We also wish to discuss a technical problem that we encountered when using the Illumina Custom Amplicon assay. Multiple point mutations with very low P-value could not be validated by Sanger sequencing, but manual review of the sequencing data (including visual examination of the alignments in Integrative Genomics Viewer [[56](#), [57](#)]) confirmed that they do exist in sequencing data. We hypothesize that this could be caused by some unknown technical artifacts introduced during the process of library preparation, where PCR may introduce errors into the sequencing data. However, on the other hand, our study used DNA extracted from FFPE tissues, yet successfully obtained sequence data on all 41 pairs of samples attempted in the study, suggesting that the protocol is quite robust to fragmented DNA molecules. Additionally, the detected somatic mutation may exist only in a small fraction of the tumor cells, which could be identified by next generation sequencing but not by Sanger sequencing. Additionally, there is also the possibility that the tumor cells mixed with stroma tissue, accounting for the difficult detection of novel recurrent somatic mutations in this study. Given that microdissection may be a better way for tumor and stroma tissue separation, we may explore it in future studies.

In conclusion, we identified recurrent mutations in genes such as *SMAD4*, *APC*, *FBXW7*, *BRAF* and *PTEN*, as well as previously unreported point mutations in gene *APC* and *TCF7L2* in a group of Chinese CRC patients. Our study represents a pilot effort to assess the effects of functional variants on four candidate pathways on CRC progression.

Supporting Information

S1 Table. Coverage description.

(XLSX)

S2 Table. Candidate somatic mutations with P-value lower than 5×10^{-5} .

(XLSX)

S3 Table. New somatic mutations detected by VarScan2.

(XLSX)

S4 Table. Reported somatic mutation in COSMIC v67 database.

(XLSX)

S5 Table. Gene exons from the array design.

(XLSX)

S6 Table. Primers used for Sanger sequencing.

(XLSX)

Acknowledgments

We thank the patients and their family members for donating tumor and adjacent stroma tissues for cancer genetic studies.

Author Contributions

Conceived and designed the experiments: KW. Performed the experiments: CL HY. Analyzed the data: CL. Contributed reagents/materials/analysis tools: CL LW ZW LX LS WDL. Wrote the paper: CL KW LW.

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, et al. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 127: 2893–2917. doi: [10.1002/ijc.25516](https://doi.org/10.1002/ijc.25516) PMID: [21351269](https://pubmed.ncbi.nlm.nih.gov/21351269/)
2. Ewing I, Hurley JJ, Josephides E, Millar A (2014) The molecular genetics of colorectal cancer. *Frontline Gastroenterol* 5: 26–30. doi: [10.1136/flgastro-2013-100329](https://doi.org/10.1136/flgastro-2013-100329) PMID: [24416503](https://pubmed.ncbi.nlm.nih.gov/24416503/)
3. Boman BM, Fields JZ (2013) An APC:WNT Counter-Current-Like Mechanism Regulates Cell Division Along the Human Colonic Crypt Axis: A Mechanism That Explains How Mutations Induce Proliferative Abnormalities That Drive Colon Cancer Development. *Front Oncol* 3: 244. doi: [10.3389/fonc.2013.00244](https://doi.org/10.3389/fonc.2013.00244) PMID: [24224156](https://pubmed.ncbi.nlm.nih.gov/24224156/)
4. Markowitz SD, Bertagnoli MM (2009) Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* 361: 2449–2460. doi: [10.1056/NEJMra0804588](https://doi.org/10.1056/NEJMra0804588) PMID: [20018966](https://pubmed.ncbi.nlm.nih.gov/20018966/)
5. Naguib A, Mitrou PN, Gay LJ, Cooke JC, Luben RN, et al. (2010) Dietary, lifestyle and clinicopathological factors associated with BRAF and K-ras mutations arising in distinct subsets of colorectal cancers in the EPIC Norfolk study. *BMC Cancer* 10: 99. doi: [10.1186/1471-2407-10-99](https://doi.org/10.1186/1471-2407-10-99) PMID: [20233436](https://pubmed.ncbi.nlm.nih.gov/20233436/)
6. Planck M, Halvarsson B, Palsson E, Hallen M, Ekelund M, et al. (2002) Cytogenetic aberrations and heterogeneity of mutations in repeat-containing genes in a colon carcinoma from a patient with hereditary nonpolyposis colorectal cancer. *Cancer Genet Cytogenet* 134: 46–54. doi: [10.1016/S0165-4608\(01\)00587-8](https://doi.org/10.1016/S0165-4608(01)00587-8) PMID: [11996796](https://pubmed.ncbi.nlm.nih.gov/11996796/)
7. Tomlinson IP, Webb E, Carvajal-Carmona L, Broderick P, Howarth K, et al. (2008) A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet* 40: 623–630. doi: [10.1038/ng.111](https://doi.org/10.1038/ng.111) PMID: [18372905](https://pubmed.ncbi.nlm.nih.gov/18372905/)
8. Tenesa A, Farrington SM, Prendergast JG, Porteous ME, Walker M, et al. (2008) Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet* 40: 631–637. doi: [10.1038/ng.133](https://doi.org/10.1038/ng.133) PMID: [18372901](https://pubmed.ncbi.nlm.nih.gov/18372901/)
9. Hutter CM, Slattery ML, Duggan DJ, Muehling J, Curtin K, et al. (2010) Characterization of the association between 8q24 and colon cancer: gene-environment exploration and meta-analysis. *BMC Cancer* 10: 670. doi: [10.1186/1471-2407-10-670](https://doi.org/10.1186/1471-2407-10-670) PMID: [21129217](https://pubmed.ncbi.nlm.nih.gov/21129217/)
10. Jia WH, Zhang B, Matsuo K, Shin A, Xiang YB, et al. (2013) Genome-wide association analyses in East Asians identify new susceptibility loci for colorectal cancer. *Nat Genet* 45: 191–196. doi: [10.1038/ng.2505](https://doi.org/10.1038/ng.2505) PMID: [23263487](https://pubmed.ncbi.nlm.nih.gov/23263487/)
11. Houlston RS, Cheadle J, Dobbins SE, Tenesa A, Jones AM, et al. (2010) Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat Genet* 42: 973–977. doi: [10.1038/ng.670](https://doi.org/10.1038/ng.670) PMID: [20972440](https://pubmed.ncbi.nlm.nih.gov/20972440/)
12. Peters U, Hutter CM, Hsu L, Schumacher FR, Conti DV, et al. (2012) Meta-analysis of new genome-wide association studies of colorectal cancer risk. *Hum Genet* 131: 217–234. doi: [10.1007/s00439-011-1055-0](https://doi.org/10.1007/s00439-011-1055-0) PMID: [21761138](https://pubmed.ncbi.nlm.nih.gov/21761138/)
13. Kapitanovic S, Cacev T, Radosevic S, Spaventi S, Spaventi R, et al. (2004) APC gene loss of heterozygosity, mutations, E1317Q, and I1307K germ-line variants in sporadic colon cancer in Croatia. *Exp Mol Pathol* 77: 193–200. doi: [10.1016/j.yexmp.2004.06.001](https://doi.org/10.1016/j.yexmp.2004.06.001) PMID: [15507235](https://pubmed.ncbi.nlm.nih.gov/15507235/)
14. Baba S, Ando H, Nakamura Y (1994) Identification of germ line mutation of APC gene in possible carriers of familial adenomatous polyposis (FAP). *Anticancer Res* 14: 2189–2192. PMID: [7840521](https://pubmed.ncbi.nlm.nih.gov/7840521/)
15. Wahlberg SS, Schmeits J, Thomas G, Loda M, Garber J, et al. (2002) Evaluation of microsatellite instability and immunohistochemistry for the prediction of germ-line MSH2 and MLH1 mutations in hereditary nonpolyposis colon cancer families. *Cancer Res* 62: 3485–3492. PMID: [12067992](https://pubmed.ncbi.nlm.nih.gov/12067992/)
16. Lynch HT, de la Chapelle A (2003) Hereditary colorectal cancer. *N Engl J Med* 348: 919–932. doi: [10.1056/NEJMra012242](https://doi.org/10.1056/NEJMra012242) PMID: [12621137](https://pubmed.ncbi.nlm.nih.gov/12621137/)

17. Casey G, Conti D, Haile R, Duggan D (2011) Next generation sequencing and a new era of medicine. *Gut* 62: 920–932. doi: [10.1136/gutjnl-2011-301935](https://doi.org/10.1136/gutjnl-2011-301935)
18. Smith G, Carey FA, Beattie J, Wilkie MJ, Lightfoot TJ, et al. (2002) Mutations in APC, Kirsten-ras, and p53—alternative genetic pathways to colorectal cancer. *Proc Natl Acad Sci U S A* 99: 9433–9438. doi: [10.1073/pnas.122612899](https://doi.org/10.1073/pnas.122612899) PMID: [12093899](https://pubmed.ncbi.nlm.nih.gov/12093899/)
19. Kinzler KW, Vogelstein B (1996) Lessons from hereditary colorectal cancer. *Cell* 87: 159–170. doi: [10.1016/S0092-8674\(00\)81333-1](https://doi.org/10.1016/S0092-8674(00)81333-1) PMID: [8861899](https://pubmed.ncbi.nlm.nih.gov/8861899/)
20. Yin Y, Tainsky MA, Bischoff FZ, Strong LC, Wahl GM (1992) Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. *Cell* 70: 937–948. doi: [10.1016/0092-8674\(92\)90244-7](https://doi.org/10.1016/0092-8674(92)90244-7) PMID: [1525830](https://pubmed.ncbi.nlm.nih.gov/1525830/)
21. Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, et al. (1991) Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* 253: 665–669. doi: [10.1126/science.1651563](https://doi.org/10.1126/science.1651563) PMID: [1651563](https://pubmed.ncbi.nlm.nih.gov/1651563/)
22. Sidransky D, Tokino T, Hamilton SR, Kinzler KW, Levin B, et al. (1992) Identification of ras oncogene mutations in the stool of patients with curable colorectal tumors. *Science* 256: 102–105. doi: [10.1126/science.1566048](https://doi.org/10.1126/science.1566048) PMID: [1566048](https://pubmed.ncbi.nlm.nih.gov/1566048/)
23. Liu Y, Bodmer WF (2006) Analysis of P53 mutations and their expression in 56 colorectal cancer cell lines. *Proc Natl Acad Sci U S A* 103: 976–981. doi: [10.1073/pnas.0510146103](https://doi.org/10.1073/pnas.0510146103) PMID: [16418264](https://pubmed.ncbi.nlm.nih.gov/16418264/)
24. Fleming NI, Jorissen RN, Mouradov D, Christie M, Sakthianandeswaren A, et al. (2013) SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. *Cancer Res* 73: 725–735. doi: [10.1158/0008-5472.CAN-12-2706](https://doi.org/10.1158/0008-5472.CAN-12-2706) PMID: [23139211](https://pubmed.ncbi.nlm.nih.gov/23139211/)
25. Jemal A, Bray F, Center MM, Ferlay J, Ward E, et al. (2011) Global cancer statistics. *CA Cancer J Clin* 61: 69–90. doi: [10.3322/caac.20107](https://doi.org/10.3322/caac.20107) PMID: [21296855](https://pubmed.ncbi.nlm.nih.gov/21296855/)
26. Hu F, Li D, Wang Y, Yao X, Zhang W, et al. (2013) Novel DNA variants and mutation frequencies of hMLH1 and hMSH2 genes in colorectal cancer in the Northeast China population. *PLoS One* 8: e60233. doi: [10.1371/journal.pone.0060233](https://doi.org/10.1371/journal.pone.0060233) PMID: [23573243](https://pubmed.ncbi.nlm.nih.gov/23573243/)
27. Thomas RK, Baker AC, Debiassi RM, Winckler W, Laframboise T, et al. (2007) High-throughput oncogene mutation profiling in human cancer. *Nat Genet* 39: 347–351. doi: [10.1038/ng0407-567a](https://doi.org/10.1038/ng0407-567a) PMID: [17293865](https://pubmed.ncbi.nlm.nih.gov/17293865/)
28. Hanna MC, Go C, Roden C, Jones RT, Pochanard P, et al. (2013) Colorectal cancers from distinct ancestral populations show variations in BRAF mutation frequency. *PLoS One* 8: e74950. doi: [10.1371/journal.pone.0074950](https://doi.org/10.1371/journal.pone.0074950) PMID: [24066160](https://pubmed.ncbi.nlm.nih.gov/24066160/)
29. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, et al. (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 22: 568–576. doi: [10.1101/gr.129684.111](https://doi.org/10.1101/gr.129684.111) PMID: [22300766](https://pubmed.ncbi.nlm.nih.gov/22300766/)
30. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38: e164. doi: [10.1093/nar/gkq603](https://doi.org/10.1093/nar/gkq603) PMID: [20601685](https://pubmed.ncbi.nlm.nih.gov/20601685/)
31. Chang X, Wang K (2012) wANNOVAR: annotating genetic variants for personal genomes via the web. *J Med Genet* 49: 433–436. doi: [10.1136/jmedgenet-2012-100918](https://doi.org/10.1136/jmedgenet-2012-100918) PMID: [22717648](https://pubmed.ncbi.nlm.nih.gov/22717648/)
32. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, et al. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13: 134. doi: [10.1186/1471-2105-13-134](https://doi.org/10.1186/1471-2105-13-134) PMID: [22708584](https://pubmed.ncbi.nlm.nih.gov/22708584/)
33. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74: 5463–5467. doi: [10.1073/pnas.74.12.5463](https://doi.org/10.1073/pnas.74.12.5463) PMID: [271968](https://pubmed.ncbi.nlm.nih.gov/271968/)
34. Smith TF, Waterman MS (1981) Identification of common molecular subsequences. *J Mol Biol* 147: 195–197. doi: [10.1016/0022-2836\(81\)90087-5](https://doi.org/10.1016/0022-2836(81)90087-5) PMID: [7265238](https://pubmed.ncbi.nlm.nih.gov/7265238/)
35. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, et al. (2011) COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* 39: D945–950. doi: [10.1093/nar/gkq929](https://doi.org/10.1093/nar/gkq929) PMID: [20952405](https://pubmed.ncbi.nlm.nih.gov/20952405/)
36. Han SW, Kim HP, Shin JY, Jeong EG, Lee WC, et al. (2013) Targeted sequencing of cancer-related genes in colorectal cancer using next-generation sequencing. *PLoS One* 8: e64271. doi: [10.1371/journal.pone.0064271](https://doi.org/10.1371/journal.pone.0064271) PMID: [23700467](https://pubmed.ncbi.nlm.nih.gov/23700467/)
37. Gay LJ, Mitrou PN, Keen J, Bowman R, Naguib A, et al. (2012) Dietary, lifestyle and clinicopathological factors associated with APC mutations and promoter methylation in colorectal cancers from the EPIC-Norfolk study. *J Pathol* 228: 405–415. doi: [10.1002/path.4085](https://doi.org/10.1002/path.4085) PMID: [22864938](https://pubmed.ncbi.nlm.nih.gov/22864938/)
38. Muzny DM, Chang K, Dinh HH, Drummond JA, et al. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487: 330–337. doi: [10.1038/nature11252](https://doi.org/10.1038/nature11252)

39. Peterson LM, Kipp BR, Halling KC, Kerr SE, Smith DI, et al. (2012) Molecular characterization of endometrial cancer: a correlative study assessing microsatellite instability, MLH1 hypermethylation, DNA mismatch repair protein expression, and PTEN, PIK3CA, KRAS, and BRAF mutation analysis. *Int J Gynecol Pathol* 31: 195–205. doi: [10.1097/PGP.0b013e318231fc51](https://doi.org/10.1097/PGP.0b013e318231fc51) PMID: [22498935](https://pubmed.ncbi.nlm.nih.gov/22498935/)
40. Hou P, Liu D, Shan Y, Hu S, Studeman K, et al. (2007) Genetic alterations and their relationship in the phosphatidylinositol 3-kinase/Akt pathway in thyroid cancer. *Clin Cancer Res* 13: 1161–1170. doi: [10.1158/1078-0432.CCR-06-1125](https://doi.org/10.1158/1078-0432.CCR-06-1125) PMID: [17317825](https://pubmed.ncbi.nlm.nih.gov/17317825/)
41. Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, et al. (2006) The consensus coding sequences of human breast and colorectal cancers. *Science* 314: 268–274. doi: [10.1126/science.1133427](https://doi.org/10.1126/science.1133427) PMID: [16959974](https://pubmed.ncbi.nlm.nih.gov/16959974/)
42. Matsumura G (1990) [Leu7 (HNK-1)-positive cells in peripheral blood and natural killer cell activity in patients with atopic dermatitis]. *Nihon Hifuka Gakkai Zasshi* 100: 57–62. PMID: [1693705](https://pubmed.ncbi.nlm.nih.gov/1693705/)
43. Kimura ET, Nikiforova MN, Zhu Z, Knauf JA, Nikiforov YE, et al. (2003) High prevalence of BRAF mutations in thyroid cancer: genetic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma. *Cancer Res* 63: 1454–1457. PMID: [12670889](https://pubmed.ncbi.nlm.nih.gov/12670889/)
44. Berg M, Danielsen SA, Ahlquist T, Merok MA, Agesen TH, et al. (2010) DNA sequence profiles of the colorectal cancer critical gene set KRAS-BRAF-PIK3CA-PTEN-TP53 related to age at disease onset. *PLoS One* 5: e13978. doi: [10.1371/journal.pone.0013978](https://doi.org/10.1371/journal.pone.0013978) PMID: [21103049](https://pubmed.ncbi.nlm.nih.gov/21103049/)
45. Zhou XP, Kuismanen S, Nystrom-Lahti M, Peltomaki P, Eng C (2002) Distinct PTEN mutational spectra in hereditary non-polyposis colon cancer syndrome-related endometrial carcinomas compared to sporadic microsatellite unstable tumors. *Hum Mol Genet* 11: 445–450. doi: [10.1093/hmg/11.4.445](https://doi.org/10.1093/hmg/11.4.445) PMID: [11854177](https://pubmed.ncbi.nlm.nih.gov/11854177/)
46. Lee JW, Soung YH, Kim SY, Lee HW, Park WS, et al. (2005) PIK3CA gene is frequently mutated in breast carcinomas and hepatocellular carcinomas. *Oncogene* 24: 1477–1480. doi: [10.1038/sj.onc.1208304](https://doi.org/10.1038/sj.onc.1208304) PMID: [15608678](https://pubmed.ncbi.nlm.nih.gov/15608678/)
47. Wu G, Xing M, Mambo E, Huang X, Liu J, et al. (2005) Somatic mutation and gain of copy number of PIK3CA in human breast cancer. *Breast Cancer Res* 7: R609–616. doi: [10.1186/bcr1262](https://doi.org/10.1186/bcr1262) PMID: [16168105](https://pubmed.ncbi.nlm.nih.gov/16168105/)
48. Ogino S, Lochhead P, Giovannucci E, Meyerhardt JA, Fuchs CS, et al. (2013) Discovery of colorectal cancer PIK3CA mutation as potential predictive biomarker: power and promise of molecular pathological epidemiology. *Oncogene*.
49. Miyaki M, Konishi M, Kikuchi-Yanoshita R, Enomoto M, Igari T, et al. (1994) Characteristics of somatic mutation of the adenomatous polyposis coli gene in colorectal tumors. *Cancer Res* 54: 3011–3020. PMID: [8187091](https://pubmed.ncbi.nlm.nih.gov/8187091/)
50. Ashton-Rickardt PG, Dunlop MG, Nakamura Y, Morris RG, Purdie CA, et al. (1989) High frequency of APC loss in sporadic colorectal carcinoma due to breaks clustered in 5q21–22. *Oncogene* 4: 1169–1174. PMID: [2797819](https://pubmed.ncbi.nlm.nih.gov/2797819/)
51. Rowan AJ, Lamlum H, Ilyas M, Wheeler J, Straub J, et al. (2000) APC mutations in sporadic colorectal tumors: A mutational “hotspot” and interdependence of the “two hits”. *Proc Natl Acad Sci U S A* 97: 3352–3357. doi: [10.1073/pnas.97.7.3352](https://doi.org/10.1073/pnas.97.7.3352) PMID: [10737795](https://pubmed.ncbi.nlm.nih.gov/10737795/)
52. Yu C, Yu J, Yao X, Wu WK, Lu Y, et al. (2014) Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing. *Cell Res* 24: 701–712. doi: [10.1038/cr.2014.43](https://doi.org/10.1038/cr.2014.43) PMID: [24699064](https://pubmed.ncbi.nlm.nih.gov/24699064/)
53. Roose J, Huls G, van Beest M, Moerer P, van der Horn K, et al. (1999) Synergy between tumor suppressor APC and the beta-catenin-Tcf4 target Tcf1. *Science* 285: 1923–1926. doi: [10.1126/science.285.5435.1923](https://doi.org/10.1126/science.285.5435.1923) PMID: [10489374](https://pubmed.ncbi.nlm.nih.gov/10489374/)
54. Zhang R, Qin W, Xu GL, Zeng FF, Li CX (2012) A meta-analysis of the prevalence of somatic mutations in the hMLH1 and hMSH2 genes in colorectal cancer. *Colorectal Dis* 14: e80–89. doi: [10.1111/j.1463-1318.2011.02858.x](https://doi.org/10.1111/j.1463-1318.2011.02858.x) PMID: [21988782](https://pubmed.ncbi.nlm.nih.gov/21988782/)
55. Vasovcak P, Pavlikova K, Sedlacek Z, Skapa P, Kouda M, et al. (2011) Molecular genetic analysis of 103 sporadic colorectal tumours in Czech patients. *PLoS One* 6: e24114. doi: [10.1371/journal.pone.0024114](https://doi.org/10.1371/journal.pone.0024114) PMID: [21901162](https://pubmed.ncbi.nlm.nih.gov/21901162/)
56. Thorvaldsdottir H, Robinson JT, Mesirov JP (2012) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 14: 178–192. doi: [10.1093/bib/bbs017](https://doi.org/10.1093/bib/bbs017) PMID: [22517427](https://pubmed.ncbi.nlm.nih.gov/22517427/)
57. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, et al. (2011) Integrative genomics viewer. *Nat Biotechnol* 29: 24–26. doi: [10.1038/nbt.1754](https://doi.org/10.1038/nbt.1754) PMID: [21221095](https://pubmed.ncbi.nlm.nih.gov/21221095/)