Cancer research

Analyzing methylated DNA in colon tumor samples for translational research

Keywords
Methylated DNA, FFPE, KingFisher Apex, MagMAX, bisulfite, TrueMark MSI, Sanger sequencing

In this application note, we show:

- DNA extracted from formalin-fixed, paraffin-embedded (FFPE) colon tumor samples using the Thermo Scientific™ KingFisher™ Apex Purification System and Applied Biosystems™ MagMAX™ FFPE DNA/RNA Ultra Kit can be analyzed for microsatellite instability (MSI) using the Applied Biosystems™ TrueMark™ MSI Assay
- Sanger sequencing of bisulfite-converted FFPE DNA can provide insight into methylation state at distinct bases in an amplicon
- The fraction of methylation at a specific base can be quantified by Sanger sequencing or by digital PCR (dPCR)
- Automated DNA recovery from FFPE samples, MSI analysis, Sanger sequencing, and dPCR provide a complete workflow for translational research of methylated DNA in colon cancer

Introduction

Globally, colon cancer (also known as colorectal cancer) is one of the most common types of cancer [1]. Typically, colorectal cancer begins as small, noncancerous polyps, which then can eventually develop into cancerous tumors [2]. Different types of colon cancers can be classified based on their genome stability, commonly measured as microsatellite instability (MSI) [3]. MSI usually arises from defects in nucleotide mismatch repair (MMR) processes, leading to expansion or contraction of repeats of simple dinucleotide, trinucleotide, or short-oligonucleotide sequences that misalign during replication [4]. Deficiencies in MMR pathways also produce mutations in coding regions of genes, leading to an increase in neoantigens in the tumor that may indicate better response to immunotherapies. Therefore, identifying MSI and its causes could be beneficial in researching an appropriate therapeutic intervention.

Several genes are known to be involved in mismatch repair in humans. *MLH1* (MutL homolog 1) is one of the crucial genes for MMR of DNA, and defects are often seen in colorectal cancer [5]. Along with other proteins, MLH1 forms a complex that recognizes and repairs mismatched bases during replication. When MLH1 activity is inhibited, these mismatches are not repaired, resulting in new mutations and microsatellite instability. MLH1 activity can be compromised by mutations that inactivate the coding region, mutations in promoters and enhancers that reduce expression, or methylation changes that can inactivate the gene [5,6].

There is growing interest in understanding the role of methylated DNA in biological processes. Changes in DNA methylation patterns are usually detected on specific sequences (for a recent review focused on cancer, see reference 7). In eukaryotes, DNA methylation is usually found as 5-methylcytosine (5-meC) in CpG sites that are present in promoters and enhancers of genes. Changes in methylation status can affect DNA–protein binding, which can often result in changes in transcription of the associated gene. In some cases, it is easier to examine methylation changes at the promoter of certain genes rather than the transcript levels, since methylated DNA is usually more stable than mRNA transcripts.

Many different techniques are used to analyze methylation changes. One of the most common ways to discover changes throughout the genome is next-generation sequencing (NGS). While NGS excels at discovering novel sequences in the genome, the relatively high expense, involved workflows, and data analysis complexity make NGS impractical for many researchers. On the other hand, PCR-based assays, also used for methylated DNA analysis, are relatively inexpensive and easy to perform, but only provide presence/absence information on the base pairs being queried.

Sanger sequencing provides another attractive option for analyzing methylated DNA sequences. The workflow is relatively simple, fast, and low-cost. Analysis of electropherogram peaks is straightforward; a reference genome based on bisulfite-modified DNA is not needed. Additionally, a single reaction can give information about all positions within an amplicon, facilitating the detection of partially modified locations in the query. Finally, by analyzing peak heights, Sanger sequencing can give an idea of what fraction of a CpG site is methylated.

Here we show how Sanger sequencing can be used to query specific methylated DNA sequences in colorectal tumor samples that have also been used for MSI analysis (Figure 1). Our proprietary software previously developed for quantifying minor alleles in Sanger sequencing traces is used to quantify the amount of methylation at a site. We also show how dPCR can be used on the same samples to precisely measure methylation at a site. Recommendations are provided for designing PCR primers used for DNA methylation analysis.

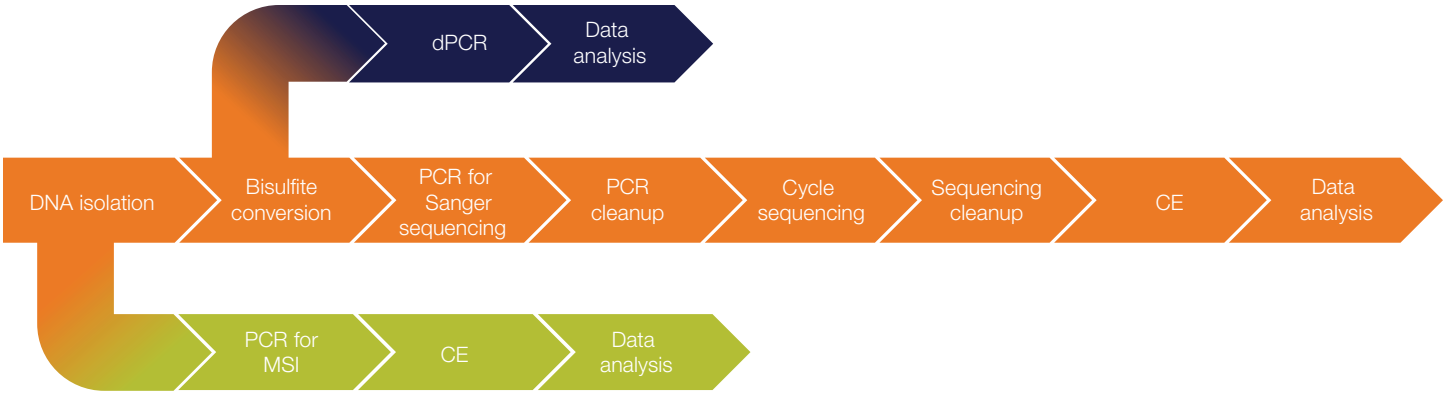


Figure 1. Workflow for analyzing methylated DNA and MSI in colon tumor samples. After extracting FFPE DNA on the KingFisher Apex instrument with the MagMAX FFPE DNA/RNA Ultra Kit, a portion of the gDNA can be used for MSI analysis using the TrueMark MSI kit and an Applied Biosystems™ Genetic Analyzer CE instrument. Another portion of gDNA can be bisulfite converted using the Thermo Scientific™ EpiJET™ Bisulfite Conversion Kit. The converted DNA can be used for Sanger sequencing with Applied Biosystems™ ExoSAP-IT™ and BigDye™ kits, or dPCR with Applied Biosystems™ QuantStudio™ Absolute Q™ products.

Sample purification on KingFisher instrument

For this study we purchased 28 FFPE colon tumor samples, and in some cases, matched normal adjacent tissue. All samples were 5 µm sections preserved on slides. Genomic DNA (gDNA) from four of the samples (CT1–CT4) and from cell lines (HT-29 and CEPH) was purified using the Invitrogen™ RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE. For the other 24 samples, gDNA was purified using the MagMAX FFPE DNA/RNA Ultra Kit on the KingFisher Apex Purification System with up-front digestion in Applied Biosystems™ AutoLys M Tubes and Caps. Briefly, each slide containing a 5 µm FFPE colon tissue section was pre-wet with 3 µL of the protease digestion buffer included in the kit. Utilizing a microtome blade, each pre-wet section was scraped off the slide and placed into its own AutoLys M Tube. The protease solution was prepared according to the MagMAX FFPE DNA/RNA Ultra Kit user guide (**Pub. No. MAN0017539**) and placed onto each section, ensuring complete submersion of the section. AutoLys M Tubes were incubated at 55°C for 3 hours followed by 90°C for 1 hour, then centrifuged for sample recovery. DNA processing plates were prepared, and the purification protocol was run on the KingFisher Apex system. Extracted DNA eluates were sealed with a foil seal and stored at –20°C until use. For the 24 samples processed on the KingFisher Apex system, we recovered up to 150 ng of gDNA (3.01 ng/µL in 50 µL). Note that in some cases the amount was below the limit of accurate detection. The concentrations of DNA recovered from the FFPE sections are summarized in Table 1.

Evaluation of MSI status

The TrueMark MSI Assay analyzes 13 microsatellite loci, including the Bethesda panel of markers [8,9], for MSI status [10]. In addition, the assay includes two highly variable short tandem repeat (STR) sequences that can be used to track sample identity. Building off an extensive database of normal samples, the analysis software does not require side-by-side analysis of normal, nontumor tissue to make MSI calls. This simplified workflow enables scientists to conserve material and quickly identify the MSI status of samples.

To determine the MSI status of the colon tumor samples, 1 µL of gDNA from the FFPE extractions, along with CEPH control and HT-29 cell line gDNA, was analyzed using the protocol and reagents provided in the TrueMark MSI kit. PCR products were run on the Applied Biosystems™ SeqStudio™ Flex Genetic Analyzer with a 50 cm capillary array and Applied Biosystems™ POP-7™ polymer. Results were analyzed using Applied Biosystems™ TrueMark™ MSI Analysis Software. MSI status was called by the software; no call (NC) status was examined manually and adjusted as needed.

Table 1. DNA recovery and microsatellite status for 28 colon tumor samples and controls.

Sample name*	DNA recovered (ng/µL)**	Tumor/ normal	Microsatellite status†
CT1	7.77	Tumor	MSS
CT2	9.57	Tumor	MSS
CT3	13.50	Tumor	MSI-L
CT4	20.20	Tumor	MSS
CT5	1.16	Tumor	MSI-L
CT6[1]	1.03	Tumor	NC (MSI-H)
CT7[1]	1.10	Normal	MSS
CT8[2]	3.01	Tumor	MSI-H
CT9[2]	0.001	Normal	NC
CT10	2.47	Tumor	MSS
CT11	1.96	Tumor	NC
CT12	0.03	Tumor	MSI-L
CT13[3]	1.20	Tumor	MSS
CT14[3]	0.93	Normal	MSS
CT15	1.32	Tumor	MSS
CT16	0.01	Tumor	MSS
CT17	0.96	Tumor	MSS
CT18	1.18	Tumor	MSS
CT19	1.20	Tumor	MSS
CT20	0.03	Tumor	NC
CT21	0.80	Tumor	MSI-H
CT22	1.23	Tumor	MSS
CT23	0.93	Tumor	MSS
CT24	0.01	Tumor	MSI-L
CT25[4]	1.51	Tumor	MSI-H
CT26[4]	1.34	Normal	MSS
CT27[5]	0.01	Normal	NC
CT28[5]	1.69	Tumor	MSS
HT-29	NA	Cell line	MSS
CEPH	NA	Normal gDNA	MSS

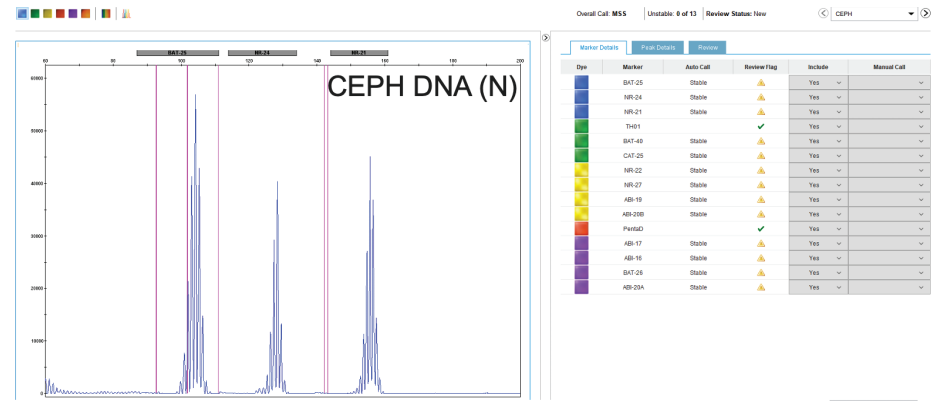
* Superscripted numbers designate tumor/normal matched samples from the same individual.

** Total eluate volume was 50 µL.

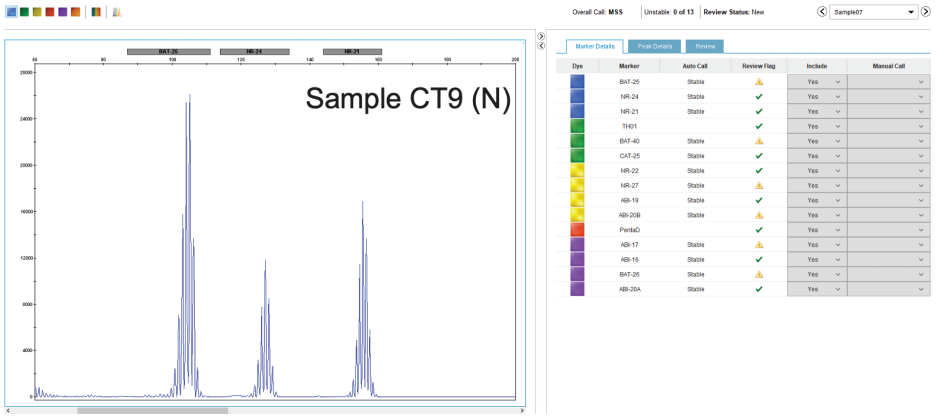
† MSS: microsatellite stable; MSI-L: low microsatellite instability (defined by software); MSI-H: high microsatellite instability; NC: no call. Sample CT6 was manually examined to determine status after a no call; other samples with NC status had no useful data upon manual examination.

The results of the microsatellite analysis are shown in Table 1. The normal CEPH gDNA was microsatellite stable (MSS), as expected (Figure 2A). In addition, the HT-29 cells were MSS, which agrees with published results [11]. Of the 5 tumor/normal matched samples, three of them had tumors that had high levels of instability (MSI-H phenotype). For example, samples CT6 and CT7 were tumor/normal matched from the same individual, and CT6 (tumor) displayed the MSI-H phenotype. CT9 (adjacent normal) showed no instability, but CT8 (tumor) clearly showed MSI at all loci (Figure 2B and C). In contrast, the other two tumors in the tumor/normal pairs (CT13 and CT28) did not show instability. Of the remaining 18 tumor samples, 2 were NC, 11 were MSS, 1 was MSI-H, and 4 had low levels of instability (MSI-L).

A



B



C

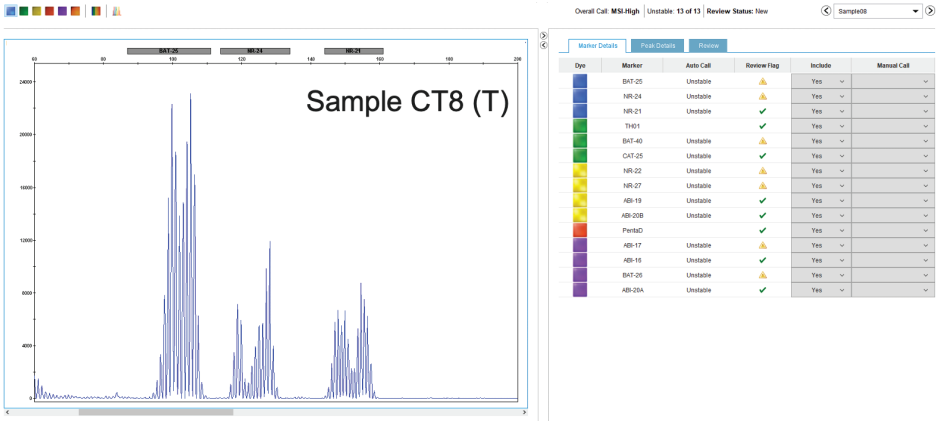


Figure 2. Examples of electropherograms and MSI calls produced by the TrueMark MSI kit. (A) CEPH DNA from normal (N) tissue shows no instability. **(B)** CT9 is a biopsy of adjacent normal tissue from the same individual as sample CT8 and has a profile similar to that of CEPH. **(C)** Sample CT8 is from a biopsy of tumor (T) tissue and shows a wider distribution of different-sized peaks than the CEPH control. Note that only the three loci in the blue channel (BAT-25, NR-24, and NR21) are shown.

Analysis of *MLH1* methylation by Sanger sequencing

Several genes are involved in mismatch repair. One of the genes commonly showing promoter methylation in colon tumors is *MLH1* [12-14]. To correlate MSI status with methylation in *MLH1*, we examined the methylation status of a portion of the *MLH1* promoter by Sanger sequencing. Sanger sequencing is ideal for focusing on a region because many potential methylation sites can be examined in a single reaction.

We analyzed methylated Cs in the *MLH1* promoter covering a 160 bp region on chromosome 3 (Table 2), as described by Benhamida et al. [13]. Primers were designed using Applied Biosystems™ Methyl Primer Express™ Software and synthesized with M13 forward or reverse tails. Extracted FFPE DNA (20 µL) was treated using the EpiJET Bisulfite Conversion Kit according to the **user guide**. Amplicons for sequencing were generated using 5 µL of bisulfite-converted DNA, the M13-tailed forward and reverse primers, and Applied Biosystems™ AmpliTaq™ Gold 360 DNA Polymerase in a total of 20 µL. Due to the reduced sequence complexity of the bisulfite-converted DNA, annealing temperatures for PCR were optimized with control DNA. Following PCR, the reactions were treated with Applied Biosystems™ ExoSAP-IT™ PCR Product Cleanup Reagent to remove excess primers. For cycle sequencing, 4 µL of the PCR product was used with the Applied Biosystems™ BigDye™ Terminator v3.1 Cycle Sequencing Kit, and either M13 forward or reverse sequencing primers. The reactions were run for 40 cycles and prepared for capillary electrophoresis (CE) with the Applied Biosystems™ BigDye XTerminator™ Purification Kit. Sequencing was performed on a SeqStudio Flex Genetic Analyzer with a 50 cm capillary array and POP-7 polymer. Data were analyzed using Geneious Prime software (**geneious.com**).

Table 2. Primers used for Sanger sequencing of bisulfite-converted DNA.

Primer name	Sequence (M13 forward and reverse tails shown in red)	Amplicon length*	Chromosome	Start (Feb 2009 GRCh37 hg19)	End (Feb 2009 GRCh37 hg19)
BCAT1 MSP F1	TGTAAAACGACGGCCAGTGTTTTTTTGTTGATGTAATTCGTTAGGTC	102	12	25,101,992	25,102,093
BCAT1 MSP R1	CAGGAAAACAGCTATGACCCAATACCCGAAACGACGACG				
FOXI2 MSP F2c	TGTAAAACGACGGCCAGTGGTTGAGCGTTAGTTTTCGTAGC	150	10	129,535,767	129,535,916
FOXI2 MSP R2bc	CAGGAAAACAGCTATGACCCCCCGAACGACCTAACGAA				
GRASP MSP F1	TGTAAAACGACGGCCAGTCGGAAGTCGCGTTCGTC	88	12	52,400,886	52,400,973
GRASP MSP R1	CAGGAAAACAGCTATGACCGCGTACAACTCGTCCGCTAA				
IKZF MSP F1	TGTAAAACGACGGCCAGTGACGACGTATTTTTTTCGTGTTTC	95	7	50,343,867	50,343,961
IKZF MSP R1	CAGGAAAACAGCTATGACCGCGCACCTCTCGACCG				
IRF4 MSP F2	TGTAAAACGACGGCCAGTTGGGTGTTTTGGACGGTTTC	110	6	392,036	392,145
IRF4 MSP R2	CAGGAAAACAGCTATGACCCGCCTACCCTCCGCG				
SOX21 MSP F3	TGTAAAACGACGGCCAGTGGCGGTGTTTTTTTTTCGC	102	13	95,364,548	95,364,649
SOX21 MSP R3	CAGGAAAACAGCTATGACCGACCGCCGCTACCTACGA				
MLH1 MSP F	TGTAAAACGACGGCCAGTTATTTTTGTTTTTATTGGTTGGAT	181	3	37,034,702	37,034,882
MLH1 MSP R	CAGGAAAACAGCTATGACCTAAATACCAATCAAATTTCTCAACTCTA				

* Amplicon length excludes M13 primers.

With Sanger sequencing, differences between unprotected, unmethylated Cs—which are converted to Ts in the sequencing trace—and protected, methylated Cs were easily detectable. In these experiments, we analyzed the alignment in the forward (F) and reverse (R) directions for each sample to ensure the calls were consistent (Figure 3). In the normal CEPH DNA sample, all the Cs in the amplicon were converted to Ts, indicating that none of the Cs were methylated. Similarly, the HT-29 cell line sample was fully unmethylated. On the other hand, one other control, a human gDNA sample that is uniformly methylated at CpGs, contained Cs at all the CpG sites, indicating that they were methylated.

We next analyzed the FFPE samples. In some of the tumor/normal pairs, differences in *MLH1* methylation were apparent. For example, the normal sample CT7 showed Ts at each possible position, indicating lack of methylation at this site (Figure 3). However, the tumor sample from the same individual (CT6) showed uniform C calls at some positions, and mixed C/T (shown as Y on the trace) at others. This suggests that methylation was indeed present at this locus in this sample. The mixed bases may reflect tumor heterogeneity in which some of the cells in the tumor had not yet acquired uniform methylation. Similarly, the normal sample CT9 showed no methylation, whereas the matched tumor (CT8) showed mixtures at each position.

Each of the tumor samples was analyzed using the same protocol (results for all samples are summarized in Figure 7). In many of the cases, there was no evidence of methylation from the resulting sequence (17 samples). In others, there was partial methylation, either across the whole region (for example, CT8, CT21, CT28), or distributed differentially across the 11 possible sites (CT6, CT15). Notably, some of those that were methylated were also MSI-H (CT6, CT8, CT21). Interestingly, one that was fully methylated was not MSI-H (CT23), and one that was MSI-H was not methylated at the *MLH1* promoter locus (CT25).

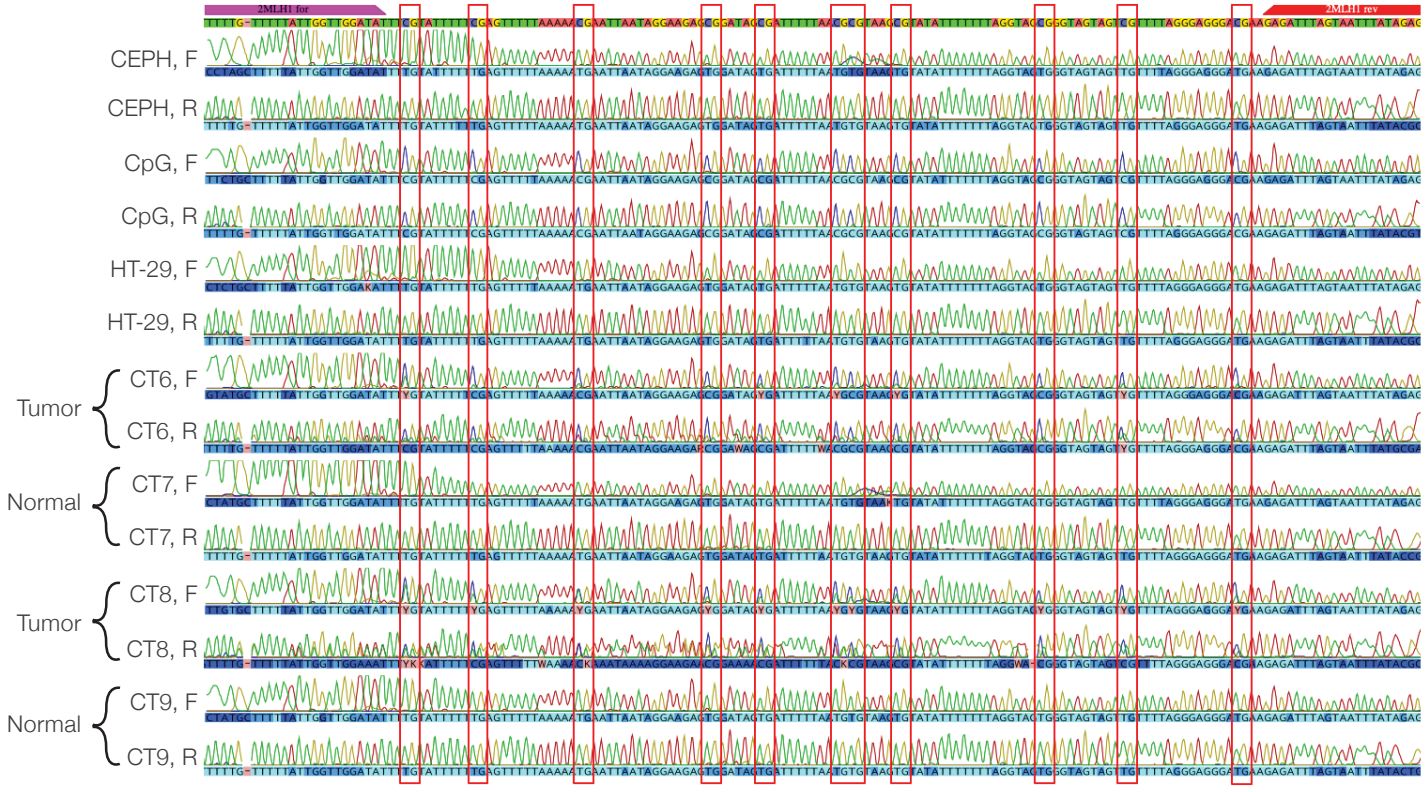


Figure 3. Sanger sequencing results of bisulfite-converted DNA at the *MLH1* locus. The positions of the sequencing primers and the bisulfite-converted sequence, with potential methylated CpGs unconverted, are shown at the top. In normal CEPH DNA, all of the Cs, including the CpGs, are converted to Ts, indicating no methylation at 11 sites (red boxes). In uniformly methylated CpG control DNA, all of the CpGs remain C, indicating they were protected from bisulfite and thus methylated. HT-29 cell line DNA is unmethylated at all CpGs. CT6 is a colon tumor sample that shows full or partial methylation at the CpGs, whereas CT7 is adjacent normal tissue from the same individual and shows no methylation. Similarly, a different tumor sample (CT8) shows there is some methylation at CpGs, whereas CT9 (matched normal) is unmethylated. Partial methylation is indicated with the mixed base call of Y (mixture of C and T at that position) and may reflect tumor heterogeneity.

Analysis of *MLH1* methylation by dPCR

The results described above suggest that in some of the samples, there was a mixture of methylated and unmethylated sequences. To confirm that the mixed-base calls were accurate, and to quantify the amount of heterogeneity, we used dPCR. We designed a custom Applied Biosystems™ TaqMan™ genotyping assay that would query the second CpG in the sequenced amplicon (Figure 4A). The probe labeled with Applied Biosystems™ FAM™ dye was designed to detect T (unmethylated at that base), and the probe labeled with Applied Biosystems™ VIC™ dye was designed to detect C (methylated). dPCR was performed using Applied Biosystems™ Absolute Q™ 1-Step RT-dPCR Master Mix (lacks uracil-DNA glycosylase), 3 µL of bisulfite-converted DNA, and a custom TaqMan genotyping assay. The reaction mix was loaded onto an Applied Biosystems™ QuantStudio™ Absolute Q™ MAP16 Plate and run on the Applied Biosystems™ QuantStudio™ Absolute Q™ Digital PCR System (96°C for 10 min precycle, then 96°C for 5 sec and 60°C for 15 sec, for 40 cycles). Data were analyzed using Applied Biosystems™ Absolute Q™ software.

The results showed clear discrimination between FAM and VIC dye signals (Figure 4B). A control sample that was fully methylated (CpG) had a very high fraction of fully methylated sequence (96.4%, Figure 4B and Table 3). Conversely, the sample that completely lacked methylation (0 Me) had very low levels of methylation (0.2%). Furthermore, methylation differences could be detected in the tumor/normal pairs. For example, the CT7 normal sample had 0.8% methylation, while the matched CT6 (tumor) had 57.2% methylation. Similarly, the matched tumor/normal pair of CT8/CT9 showed increased methylation in the tumor sample (31.3% vs. 0.3%). Importantly, these values correlate with the mixed peaks detected in the Sanger sequencing traces described above.

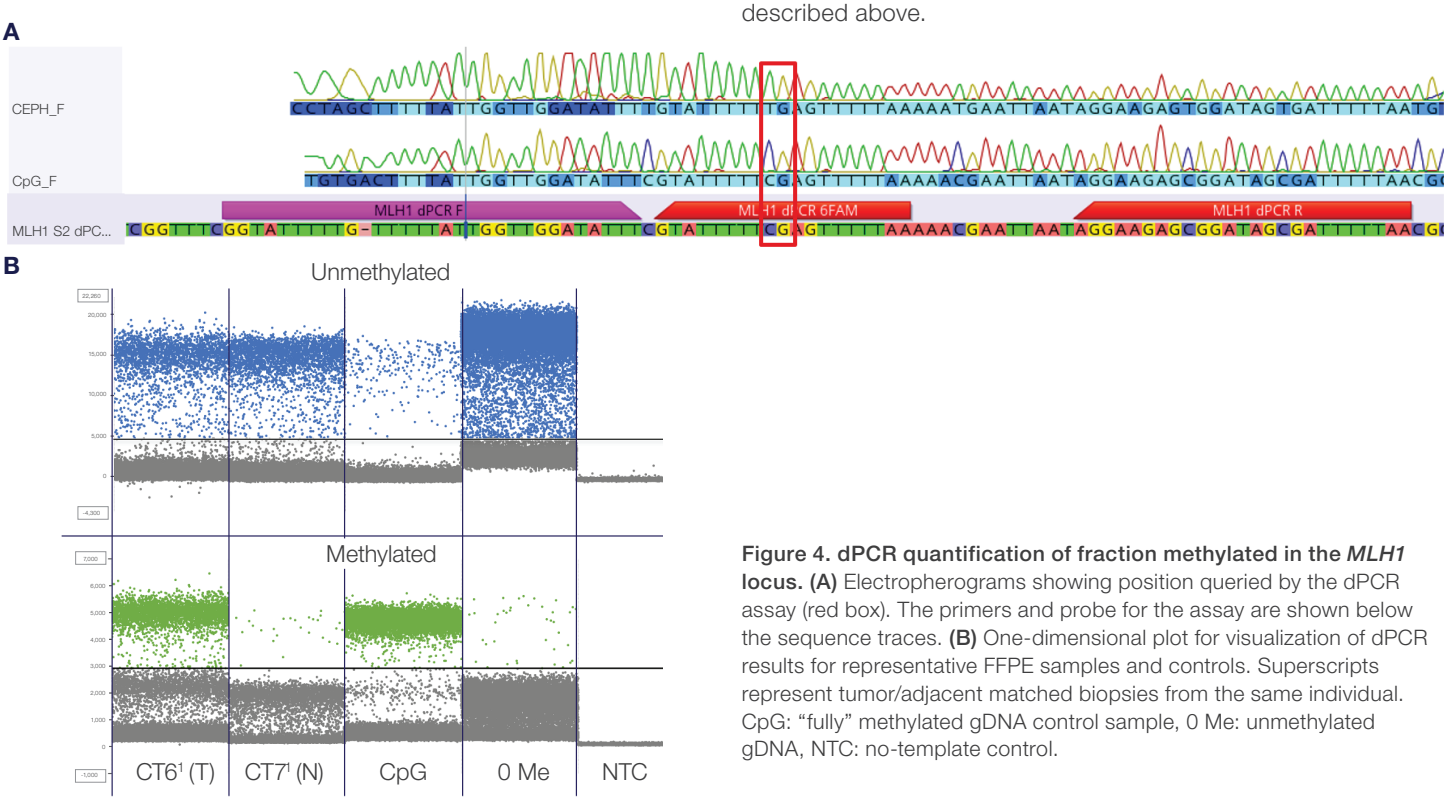


Figure 4. dPCR quantification of fraction methylated in the *MLH1* locus. (A) Electropherograms showing position queried by the dPCR assay (red box). The primers and probe for the assay are shown below the sequence traces. **(B)** One-dimensional plot for visualization of dPCR results for representative FFPE samples and controls. Superscripts represent tumor/adjacent matched biopsies from the same individual. CpG: "fully" methylated gDNA control sample, 0 Me: unmethylated gDNA, NTC: no-template control.

Table 3. Fraction of methylation determined by dPCR for the indicated samples.

	FAM channel (unmethylated, copies/µL)	VIC channel (methylated, copies/µL)	Percent methylated*
CT6[1]	309.17	413.45	57.2
CT7[1]	376.94	3.06	0.8
CT8[2]	883.65	403.21	31.3
CT9[2]	551.33	1.58	0.3
CT21	1,083.99	967.55	47.2
CpG	31.43	841.30	96.4
FMG**	237.18	5,390.22	95.8
0 Me	2,900.82	4.53	0.2

* Calculated by dividing copies/µL in the VIC channel by the sum of copies/µL in the FAM and VIC channels.
** Fully methylated sample from a different source.

Analysis of methylation at other loci

Mitchell et al. [15] described a panel of loci that had a high degree of methylation in colorectal cancer, and thus all or some of these loci may be useful as biomarkers for colorectal cancers. To determine whether these markers would be useful for analysis by Sanger sequencing, we synthesized a subset of the PCR primers described in that reference, adding M13 tails for subsequent sequencing (Table 2). For sequencing, we followed the protocol described above for *MLH1*, using 1 µL (instead of 5 µL) of bisulfite-converted DNA.

Although not all samples produced results, methylated CpGs were clearly detectable in those that were successful (not all sequencing traces are shown here). For example, the *GRASP1* amplicon showed clear methylation in the first four CpGs, but partial (CT17) or no methylation (CT16) at the fifth CpG (Figure 5). Similarly, methylation differences could be seen in the third, fourth, and fifth CpG of *IRF4* in samples CT6, CT10, and CT12, while the first, second, and sixth were uniformly methylated (Figure 6). A summary of all the Sanger sequencing results is shown in Figure 7.

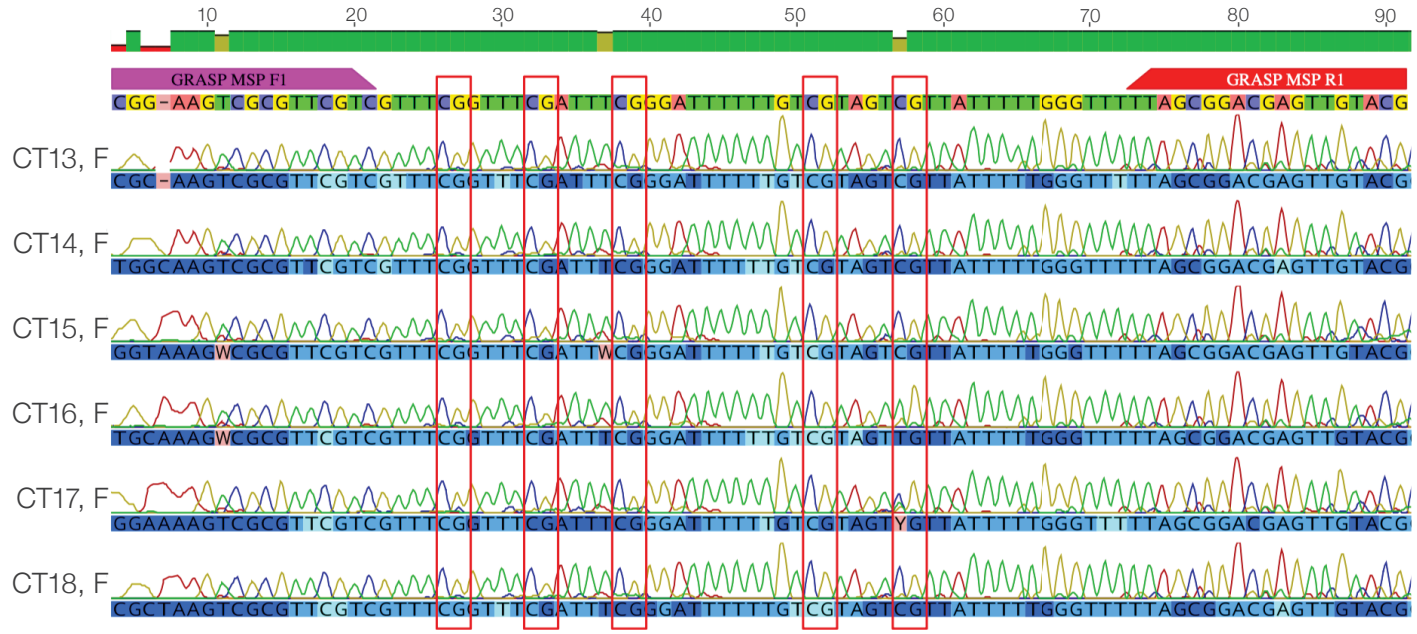


Figure 5. Sanger sequencing results of bisulfite-converted DNA at the *GRASP1* locus from six colon tumor samples. The positions of the sequencing primers and the bisulfite-converted sequence, with potential methylated CpGs unconverted, are shown at the top. Forward-direction sequencing traces of six representative tumor samples are shown. Red boxes indicate queried CpGs. Note that although most are methylated, in one case methylation is lacking (CT16) or is heterogeneous (CT17) at the fifth CpG in the amplicon.

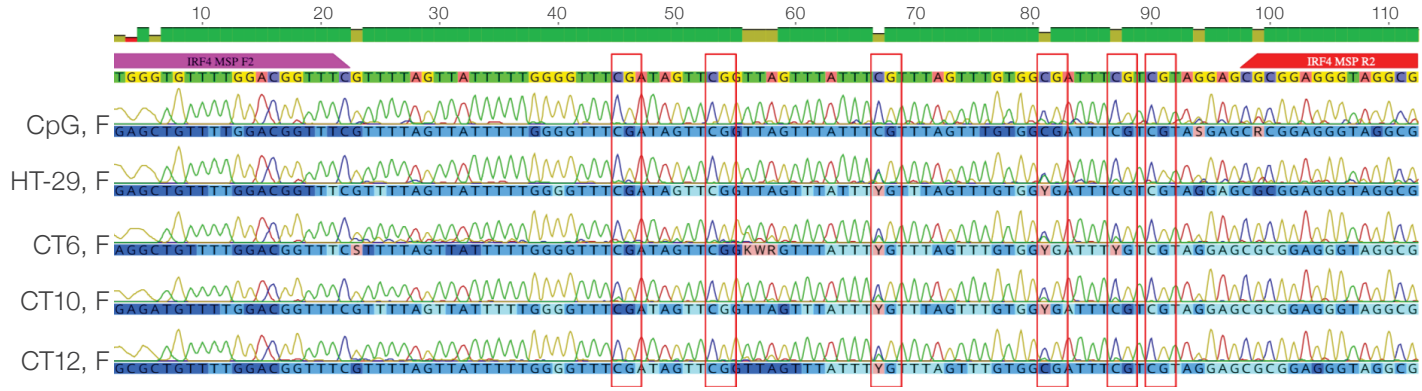


Figure 6. Sanger sequencing results of bisulfite-converted DNA at the *IRF4* locus from three colon tumor samples and two controls. The positions of the sequencing primers and the bisulfite-converted sequence, with potential methylated CpGs unconverted, are shown at the top. Forward-direction sequencing traces of three representative tumor samples are shown. Red boxes indicate queried CpGs. Note that the first, second, and sixth CpG are almost uniformly methylated, but the third and fourth CpGs show mixed bases and thus have mixed status.

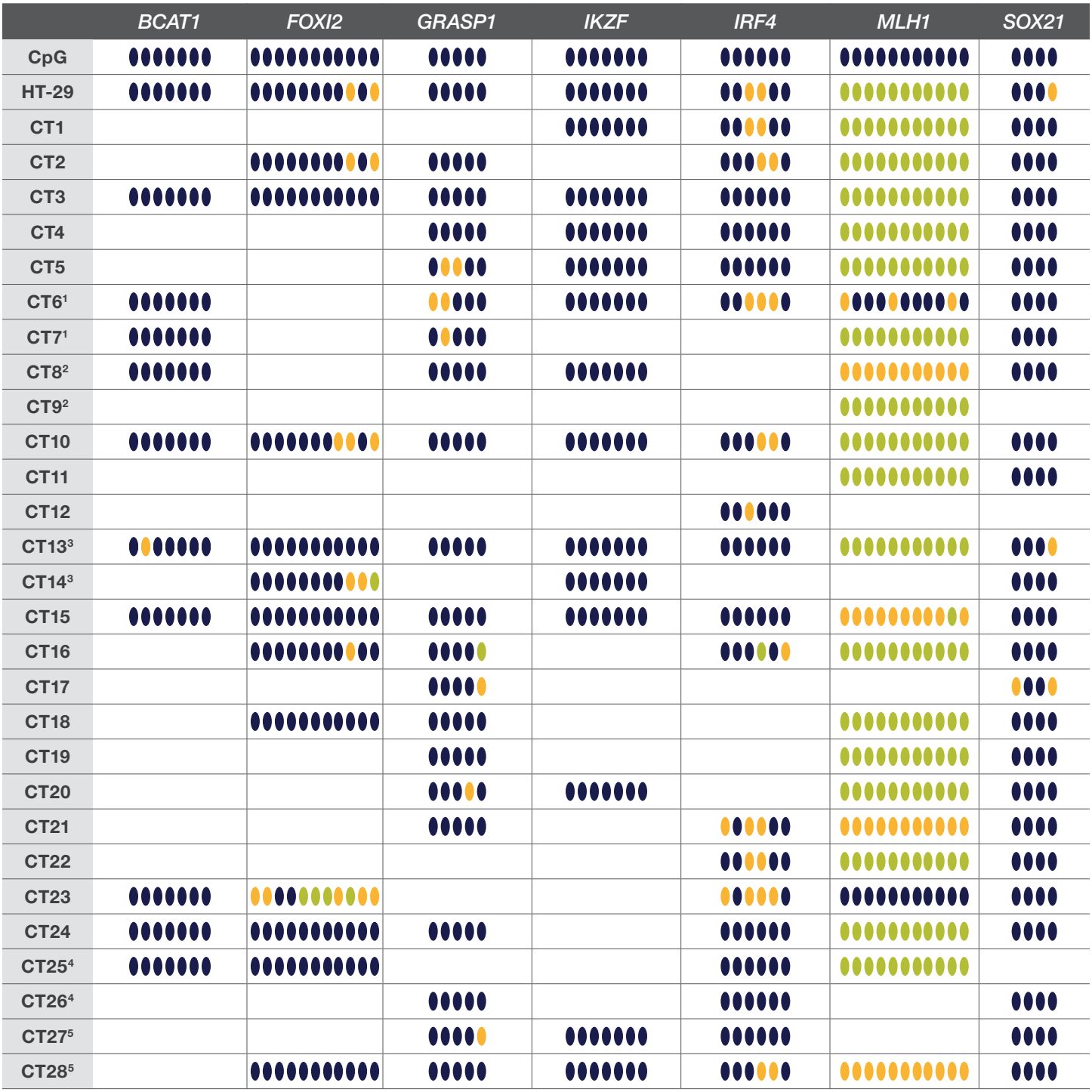


Figure 7. Summary of methylation determined by Sanger sequencing. Each CpG site in a locus is represented by an oval. Dark blue oval: fully methylated at that base; green oval: fully unmethylated at that base; orange oval: partially methylated at that base (mixed base call). Superscripts designate tumor/normal adjacent pairs from the same individual. Blank spaces: no useful sequence was obtained and material was lacking in repeats.

Applied Biosystems™ Minor Variant Finder Software (thermofisher.com/mvf) was developed to facilitate the analysis of allelic mixtures using Sanger sequencing traces. This free software compares the areas contributed by the different alleles under a mixed peak, and after subtracting background, gives the relative contributions of the different alleles. We tested whether Minor Variant Finder Software could be used to quantify partial methylation in the *IRF4* amplicon, which shows mixed bases at the third CpG in several samples. Using the CpG gDNA as the reference sample (fully methylated, so C is the reference base and T is the minor allele), Minor Variant Finder Software analysis of CT10, HT-29, and CT16 traces showed that the fraction methylated at position 80 of the amplicon (the third CpG) was 83.4%, 62.2%, and 20.9% (average of quantification in forward and reverse directions) in CT10, HT-29, and CT16 samples, respectively (Figure 8). Note: Since the minor allele detected by Minor Variant Finder Software in this case is unmethylated, we subtracted the measurement from 100 to obtain the fraction that was methylated.

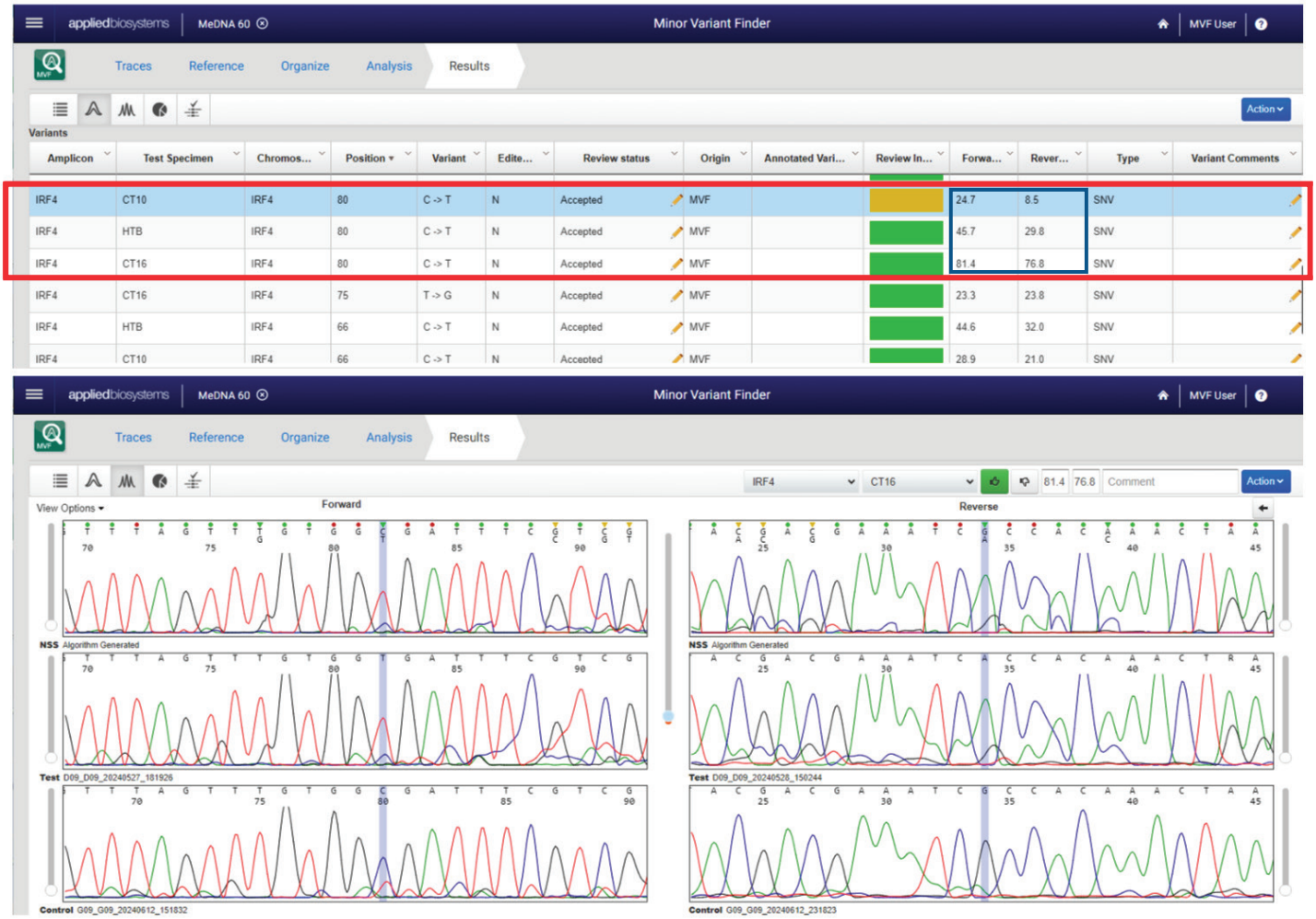


Figure 8. Quantification of fraction methylated at one position in the *IRF4* locus. Sequencing traces from two colon tumor samples and the cell line HT-29 were analyzed using Minor Variant Finder Software. The CpG traces were used as the "normal" reference. Here, the "normal" base at position 80 (third CpG) is methylated C, while unmethylated C (T after bisulfite conversion) is the minor allele. Results of the quantification are shown at the top (red box); values for each sample in the forward and reverse directions (blue box) are averaged, and then that average is subtracted from 100 to calculate percent methylation. Traces for sample CT16 and the CpG control are shown at the bottom.

We used dPCR on the QuantStudio Absolute Q system to confirm these results. We synthesized a custom TaqMan genotyping assay to query the third CpG in *IRF4* (Figure 9). dPCR was performed as described above for *MLH1*, except using 1 µL of bisulfite-converted DNA. The amount of methylation could be subsequently quantified by comparing the quantities measured in the FAM channel (unmethylated) to those in the VIC channel (methylated). The CpG fully methylated control showed 97.65% of the base was methylated in the sample (Table 4). Results for a representation of tumor samples are shown as well. Note that for samples CT10 and HT-29, we measured 67.01% and 60.74% methylation, respectively. These values were similar to the measurements obtained by Minor Variant Finder Software and Sanger sequencing (83.4% and 62.2%, respectively).

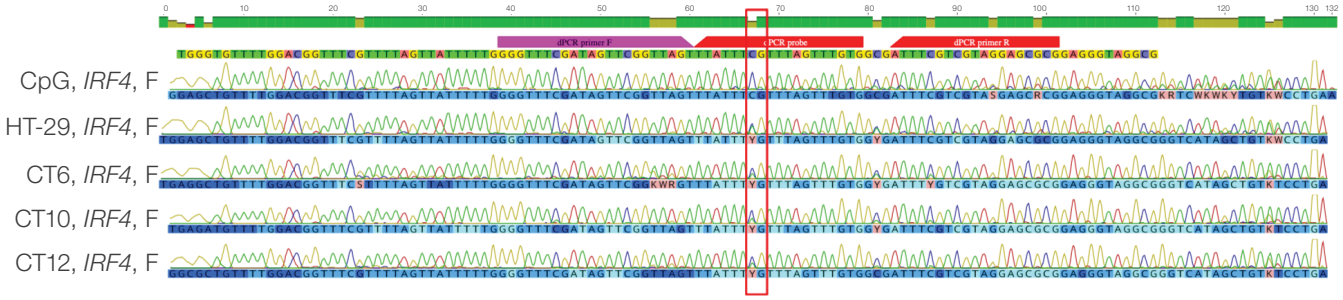


Figure 9. Quantification of methylation at the *IRF4* locus. Electropherograms showing position of bases (red box) that were queried by the dPCR primers and probe.

Table 4. Fraction of methylated Cs at the position shown in Figure 9, as determined by dPCR.

	FAM channel (unmethylated; copies/µL)	VIC channel (methylated; copies/µL)	Percent methylated*
HT-29	31.50	48.73	60.74
CT6	151.01	176.45	53.88
CT23	10.79	9.06	45.64
CT10	80.75	164.05	67.01
CT2	23.92	58.87	71.11
CT12	2.95	29.24	90.84
CT22	19.57	139.12	87.67
CpG	1.13	47.01	97.65

* Calculated by dividing copies/µL in the VIC channel by the sum of copies/µL in the FAM and VIC channels.

Summary

Here we have demonstrated how methylated DNA can be analyzed in a workflow optimized for translational research. We show how high-quality gDNA extracted from small FFPE sections using the MagMAX FFPE DNA/RNA Ultra Kit on the KingFisher Apex system can be analyzed using the TrueMark MSI Assay. We described how the extracted DNA can be used to analyze methylation status in specific regions using Sanger sequencing. The amount of methylation at specific bases can be quantified using Sanger sequencing and Minor Variant Finder Software, or using TaqMan genotyping assays on the QuantStudio Absolute Q Digital PCR System.

Sanger sequencing is widely used to analyze sequences because of its ease of use, simple data analysis, and relatively low cost. These features also make it ideal for analyzing methylated DNA across defined regions. Unlike PCR-based approaches, Sanger sequencing can provide information about the methylation state at many bases in an amplicon. And unlike NGS-based approaches, the resulting electropherograms are easily interpreted and do not require separate reference genomes for analyzing the different methylation states that are possible. Sanger sequencing is therefore useful for confirming methylation states over defined regions and is ideal for confirming discovery-based results, for biomarker analysis, and for linking epigenetic results to biological outcomes and pathways.

We show that the fraction that is methylated in a sample can be quantified using Sanger sequencing and Minor Variant Finder Software, or using custom TaqMan Assays on the QuantStudio Absolute Q Digital PCR System. The fractional methylation we observed may have been due to tumor heterogeneity or might be important in an underlying biological process. Using Minor Variant Finder Software and Sanger sequencing, differences down to 5% abundance can be detected [16], while dPCR can detect 0.1% or lower abundance. None of the samples analyzed in this application note had methylation differences at those low frequencies. In general, the choice of system to use for methylated DNA analysis depends on the research goals. For simultaneous analysis of multiple CpG sites, Sanger sequencing is likely to be the best choice. For projects that require precise quantification of methylation at single CpG sites, dPCR may be more appropriate.

The analysis of methylated DNA depended on using bisulfite-converted DNA, resulting in some challenges. First, the process of bisulfite conversion fragments DNA, so for best results the amplicons to be analyzed should be 100–200 nucleotides long. Second, because bisulfite conversion changes unmethylated Cs to Ts, it reduces the complexity of the genomic sequence, resulting in fewer possibilities for optimal primers. Finally, because any primers will likely be relatively AT-rich, melting temperatures may be lower than for unmodified DNA. In some cases, it might not be possible to find optimal primers. We recommend optimizing primer design and PCR conditions, including annealing temperature and time, before proceeding with precious samples.

The analysis of methylated DNA has enormous potential for understanding gene regulation, epigenetic inheritance, and mechanisms of some pathologies, and for biomarker discovery and deployment. The workflows described above will give researchers tools that will help realize those possibilities. For more information, please fill out this **contact form** and a representative will be in touch.

Ordering information

Product	Cat. No.
MagMAX FFPE DNA/RNA Ultra Kit	A31881
AutoLys M Tubes and Caps	A38738
AutoLys M Tube Racks	A37955
Fisherbrand Isotemp General Purpose Heating and Drying Oven	15-103-0510
TrueMark MSI Assay	A45295
AmpliTaq Gold 360 Master Mix	4398876
ExoSAP-IT Express PCR Product Cleanup	75001.40.UL
BigDye Terminator v3.1 Cycle Sequencing Kit	4337454
BigDye XTerminator Purification Kit	4376486
Custom TaqMan SNP Genotyping Assay, nonhuman	4332076
Absolute Q 1-Step RT-dPCR Master Mix (4X)	A55146
QuantStudio Absolute Q MAP16 Plate Kit	A52865
CpG Methylated Human Genomic DNA	SD1131
KingFisher Apex Purification System with 96 Deep-Well Head	5400930
ProFlex 96-Well PCR System	4484075
SeqStudio 8 Flex Genetic Analyzer	A53627
QuantStudio Absolute Q Digital PCR System, desktop	A52864
Absolute Q Universal DNA dPCR Master Mix (5X)	A72710
Minor Variant Finder Software	A30835

References

1. Bray F et al. (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68:394–424. **doi.org/10.3322/caac.21492**
2. Sullivan BA et al. (2022) Cause, epidemiology, and histology of polyps and pathways to colorectal cancer. *Gastrointest Endosc Clin N Am* 32(2):177–194. **doi.org/10.1016/j.giec.2021.12.001**
3. Zeinalian M et al. (2018) Clinical aspects of microsatellite instability testing in colorectal cancer. *Adv Biomed Res* 7:28. **doi.org/10.4103/abr.abr_185_16**
4. Reyes GX et al. (2015) New insights into the mechanism of DNA mismatch repair. *Chromosoma* 124:443–462. **doi.org/10.1007/s00412-015-0514-0**
5. Tariq K et al. (2016) Colorectal cancer carcinogenesis: a review of mechanisms. *Cancer Biol Med* 13(1):120–135.
6. Mitchell RJ et al. (2002) Mismatch repair genes *hMLH1* and *hMSH2* and colorectal cancer: A HuGE review. *Am J Epidemiol* 156(10)885–902. **doi.org/10.1093/aje/kwf139**
7. Papanicolau-Sengos A et al. (2022) DNA methylation profiling: An emerging paradigm for cancer diagnosis. *Annu Rev Pathol* 17:295–321. **doi.org/10.1146/annurev-pathol-042220-022304**
8. Vilar E et al. (2010) Microsatellite instability in colorectal cancer—the stable evidence. *Nat Rev Clin Oncol* 7:153–162. **doi.org/10.1038/nrclinonc.2009.237**
9. Boland CR et al. (1998) A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 58:5248–5257.
10. TrueMark MSI research assay for microsatellite instability analysis. **thermofisher.com/us/en/home/clinical/clinical-genomics/molecular-oncology-solutions/microsatellite-instability-assay.html**
11. Duldulao MP et al. (2012) Gene expression variations in microsatellite stable and unstable colon cancer cells. *J Surg Res* 174(1):1–6. **doi.org/10.1016/j.jss.2011.06.016**
12. Morak M et al. (2018) Comprehensive analysis of the *MLH1* promoter region in 480 patients with colorectal cancer in 1150 controls reveals new variants including one with a heritable constitutional *MHL1* epimutation. *J Med Genet* 55:240–248. **doi.org/10.1136/jmedgenet-2017-104744**
13. Benhamida JK et al. (2020) Reliable clinical *MLH1* promoter hypermethylation assessment using a high-throughput genome-wide methylation array platform. *J Mol Diag* 22(3):368–375. **doi.org/10.1016/j.jmoldx.2019.11.005**
14. Hitchens MP et al. (2023) Constitutional *MLH1* methylation is a major contributor to mismatch repair deficient, *MLH1*-methylated colorectal cancer in patients aged 55 years and younger. *J Natl Compr Canc Netw* 21(7):743–752.e11. **doi.org/10.6004/jnccn.2023.7020**
15. Mitchell SM et al. (2014) A panel of genes methylated with high frequency in colorectal cancer. *BMC Cancer* 14:54. **doi.org/10.1186/1471-2407-14-54**
16. Minor Variant Finder Software. **thermofisher.com/mvf**

Learn more at **thermofisher.com/abcancerresearch**

applied biosystems