

Development of a Liquid Biopsy Based Purely Quantitative Digital Droplet PCR Assay for Detection of MLH1 Promoter Methylation in Colorectal Cancer Patients

Danyi Wang

EMD Serono Research and Development Institute

Dennis O'Rourke

EMD Serono Research and Development Institute

Jorge Sanchez Garcia

EMD Serono Research and Development Institute

Ti Cai

EMD Serono Research and Development Institute

Juergen Scheuenpflug

Merck Biopharma; Merck KGaA

Zheng Feng (✉ zheng.feng@emdserono.com)

EMD Serono Research and Development Institute

Research Article

Keywords: digital droplet PCR (ddPCR) assay, MLH1 promoter methylation, colorectal cancer

Posted Date: March 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-155375/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Cancer on July 10th, 2021. See the published version at <https://doi.org/10.1186/s12885-021-08497-x>.

1 **Development of a Liquid Biopsy Based Purely Quantitative Digital Droplet PCR Assay for Detection of**
2 ***MLH1* Promoter Methylation in Colorectal Cancer Patients**

3 Danyi Wang, PhD^{1,2#}, Dennis O'Rourke, BS^{1,2#}, Jorge F Sanchez-Garcia, PhD^{1,2}, Ti Cai, PhD^{1,2}, Juergen
4 Scheuenpflug, PhD^{1,3} and Zheng Feng, PhD, MD^{1,2*}

5

6 1. Global Clinical Biomarkers and Companion Diagnostics, Global Early Development

7 2. EMD Serono Research and Development Institute, Billerica, MA, USA

8 3. Merck Biopharma; Merck KGaA, Darmstadt, Germany

9

10 # Co-first author

11 * Corresponding author

12

13 Keywords: *MutL Homolog 1*, Circulating tumor DNA, Methylation, Methylation sensitive restriction enzyme,
14 Droplet digital PCR

15

16 **Abstract**

17 **Background**

18 *MutL Homolog 1 (MLH1)* promotor methylation is associated with microsatellite instability high colorectal

19 cancer (CRC). The strong correlation between methylation status and cancer development and progression

20 has led to a growing interest in the use of methylation markers in circulating tumor DNA (ctDNA) for early

21 cancer detection and longitudinal monitoring. As cancer-specific DNA methylation changes in body fluids are

22 limited, it is particularly challenging to develop clinically applicable liquid biopsy methodologies with high

23 sensitivity and specificity. The purpose of this study was to develop a fit-for-purpose methylation sensitive

24 restriction enzyme (MSRE) based digital droplet PCR (ddPCR) assay to examine *MLH1* promoter

25 methylation in ctDNA in advanced CRC.

26 **Methods**

27 Primers and probes were designed to amplify CpG sites of the *MLH1* promoter. Methylated and unmethylated

28 control genomic DNA were sheared to mimic ctDNA and subjected to MSRE *HpaII* digestion. Plasma samples

29 from 20 healthy donors and 28 CRC patients were analyzed with the optimized MSRE procedure using ddPCR.

30 **Results**

31 Using methylated and unmethylated controls, we optimized the conditions for *HpaII* enzyme digestion
32 to ensure complete digestion and avoid false positives. Based on the results from the ddPCR assay using
33 1ng circulating cell-free DNA (cfDNA) input from healthy donors or CRC samples, ROC curves were
34 generated with an area under the curve (AUC) value of 0.965. The optimal assay sensitivity and
35 specificity was achieved when 8 positive droplets were used as acceptance criteria (78% sensitivity and
36 100% specificity). A tiered-based cutoff (20%, 50%, 80% percentile based) was applied to distinguish CRC
37 samples with different methylation level.

38 **Conclusions**

39 Our study demonstrated that the liquid biopsy assay for *MLH1* promoter methylation detection using
40 purely quantitative ddPCR is a simple and ultrasensitive procedure that provides reliable methylation
41 detection in ctDNA. The MSRE ddPCR approach can also be applied to other genes of interest where
42 methylation patterns could reveal clinically relevant information for future clinical biomarker and/or
43 companion diagnostic development.

44

45 **Background**

46 Methylation changes are present in a variety of cancers, and occur early in carcinogenesis, typically
47 repressing the expression of tumor suppressor genes. Acquired promoter hypermethylation often
48 occurs with global hypermethylation of gene promoters known as CpG island methylator phenotype [1].
49 These epigenetic changes are highly pervasive across a tumor type and can be a very consistent feature
50 of cancer in contrast to mutations, which typically occur at a wide range of sites. *MLH1* promoter
51 hypermethylation is an important event, silencing the *MLH1* gene expression and preventing the
52 formation of MLH1 protein and normal activation of the DNA repair gene. *MLH1* promoter methylation
53 in sporadic tumors causes high levels of microsatellite instability (MSI). Approximately 15% of CRC cases

54 are MSI-High (MSI-H); 3% of which include hereditary polyposis colorectal cancer (HNPCC) or Lynch
55 syndrome and sporadic MSI-H CRC, typically caused by somatic methylation of the *MLH1* gene
56 promoter, making up the remaining 12% of cases. [2, 3]. Specifically, within gastric cancer, methylation
57 of CpG islands in the *MLH1* promoter region is frequent resulting in MSI and *MLH1* inactivation. [4][5-7].
58 Emerging evidence demonstrated that detection of DNA methylation-based biomarkers could be
59 employed for diagnosis, indicating prognosis and predicting response to therapy [8-10]. Therefore, the
60 assessment of *MLH1* promoter methylation status will add essential clinical utility values across the
61 cancer types.

62 DNA methylation analysis is not only limited to tissue specimens, but also can be readily
63 extended to be detected in cfDNA/ctDNA, which enables a non-invasive clinical solution. Importantly,
64 the robust and common nature of DNA methylation aberrations in cancer, and the stability of cell-free
65 DNA in body fluids are attractive properties for biomarker development. Recently, MSI status could be
66 determined using blood ctDNA test with comparable performance as tissue biopsies-based test [11],
67 potentially providing an alternative pan cancer based approach to monitor patient's response to
68 therapies.

69 One of the critical challenges to establishing a DNA methylation assay with clinical usefulness, is
70 identifying the proper region of interest. The investigated region should ideally fulfill the following
71 criteria: first, the region should be unmethylated in normal cases and methylated only in cancer cases;
72 and second, the methylation levels of this region should clearly allow the classification of two distinct
73 populations (for example high versus low or cancer versus healthy). As methylation markers mainly
74 depend on the methylation level of individual CpG sites, the assay sensitivity is limited by the technical
75 difficulties in measuring single-CpG methylation.

76 Current approaches to methylation analysis depend on bisulfite sequencing, which may cause a
77 degree of DNA degradation and reduce assay reproducibility. Whole-genome bisulfite Sequencing

78 (WGBS) provides complete methylation information, however this approach is not cost-efficient and has
79 no clinical application path. Subsequent optimized approaches such as reduced-representation bisulfite
80 sequencing (RRBS-Seq) and methylated CpG tandems amplification and sequencing (MCTA-seq) aim at
81 capturing CpG-enriched cfDNA fragments, which would lead to loss of some critical DNA
82 methylation sites [12]. Therefore, it is essential to develop a cost-efficient ctDNA methylation analysis
83 approach with improved sensitivity and specificity which could be implemented in a clinical setting. In
84 this study, we developed a methylation sensitive restriction enzyme (MSRE) digested assay to absolutely
85 quantify the *MLH1* promoter methylation in ctDNA using highly sensitive ddPCR in advanced stage CRC
86 samples.

87

88 **Materials and Methods**

89 **Sample preparation**

90 *Preparation of sheared methylated/unmethylated DNA:* EpiScope® Unmethylated HCT116 DKO gDNA and
91 EpiScope® Methylated HCT116 gDNA reference cell-line DNA samples (catalog #3521, #3522 Takara Bio,
92 Mountainview CA) were sheared to approximately 180 bp by acoustic shearing to mimic fragmented
93 cfDNA using a Covaris Ultrasonicator (Covaris, Woburn MA). DNA size distribution of sheared DNA was
94 tested using Agilent 2100 Bioanalyzer High Sensitivity DNA Kit (catalog #5067-4626, Agilent
95 Technologies, Santa Clara CA) and quantified by Qubit dsDNA High Sensitivity kit (catalog #Q32854,
96 ThermoFisher Scientific, Waltham MA). Fully methylated and unmethylated DNA were mixed to obtain
97 the following ratios of methylation: 0%, 10%, 25%, 50%, 75%, 100%. Standard curves with known
98 methylation ratios were included in the assay to deduce the methylation level of tumor samples.

99 *Sample acquisition and cfDNA extraction:* Healthy donor (n= 20, median age 44 with range 22-71) and
100 advanced colorectal cancer samples (n= 28 (stage IIIB: 21, IIIC: 4, IV:3), median age 63 with range 32-77)

101 used in this study were procured from BioIVT. The written informed consent was received from all
102 subjects used in this study and all samples were collected under IRB approved protocols, as defined by
103 the specimen provider BioIVT. Ethical approval was obtained from BioIVT ethics committee. This study
104 was conducted under the guidelines put forth into the Declaration of Helsinki. No clinical treatment was
105 applied to the CRC patients before blood collection. Specifically, blood samples were collected using
106 blood collection tubes containing EDTA (BD Vacutainer®). Plasma isolation was performed within 4
107 hours of blood collection using a double spin protocol. Plasma nucleic acid was extracted using QIAamp
108 MinElute ccfDNA Kit (catalog #55204, Qiagen, Valencia CA) according to manufacturer's instructions.
109 Approximately 5mL EDTA plasma was used for the extraction of cfDNA from normal patient samples.
110 CRC plasma samples ranged in volume from 1-4mL and were brought up to 5mL each by adding PBS to
111 necessary samples for uniform extraction and to avoid freeze/thaw cycles. Purified nucleic acids were
112 collected in 35µl elution buffer for both normal samples and CRC samples. Two microliters were used for
113 quantification of extracted cfDNA using Qubit and Bioanalyzer.

114

115 **Methylation sensitive restriction enzyme (MSRE) digestion**

116 MSREs cleave DNA at specific unmethylated-cytosine residues and DNA is then amplified by PCR
117 following digestion. As unmethylated DNA is digested by the MSRE reaction, the only amplification
118 products detected are methylated DNA. Here, a MSRE protocol was optimized for selected *MLH1*
119 promoter region by *HpaII* enzymes. To ensure full digestion of the DNA before PCR amplification to
120 generate accurate data, different *HpaII* enzymes were tested at various concentrations, and with various
121 incubation times. First, two *HpaII* enzymes were identified and tested for complete digestion.
122 Methylated DNA standards and unmethylated DNA standards were treated with each enzyme per
123 manufacturer's recommendation as well as run in mock digestion reactions where no enzyme was
124 added. These reactions were tested in qPCR and the calculated percent methylation was determined.

125 The percent methylation was calculated by subtracting the mean Cq values of the MSRE digested
126 templates from the corresponding mock digest (ΔCq). The methylation level was then calculated using
127 the formula;

$$128 \quad \text{Methylation level} = (2^{-\Delta Cq}) * 100$$

129 Ultimately, we determined the ThermoFisher *HpaII* enzyme (catalog #ER0511 Waltham, MA
130 USA) was the better of the two enzymes (data not shown). Next, we tested different incubation times
131 (2-16 hours) and concentrations (2 μ L/reaction and 0.5 μ L/reaction) to ensure complete digest of the
132 unmethylated samples. Using the same calculations as above, we determined a 3hour incubation with
133 2 μ L of enzyme per reaction was sufficient for complete digestion and extending the incubation time did
134 not significantly improve digestion. Our final MSRE reaction contained: 2 μ l 10 \times Tango Buffer 3, 2 μ l
135 *HpaII* enzyme and 5ng DNA in total 20 μ l final volume. Mock reactions contained 2 μ l 10 \times Tango Buffer 3
136 and 2 μ l *HpaII* enzyme. Reactions were incubated at 37°C for 3 hours and subsequently heat inactivated
137 for 20 min at 65°C. The resulting, 0.25 ng/ μ L digested DNA was stored at -20°C for subsequent analysis.

138

139 **Primers and probes design**

140 Beta-Actin (*ACTB*) and Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) genes were used as the
141 endogenous reference genes for qPCR and acquired from ThermoFisher (*ACTB* Taqman Assay ID:
142 Hs01060665_g1, *GAPDH* Taqman Assay ID: Hs02786624_g1, FAM-MGB). Primers and probe (Forward:
143 5'-GATGAGGCGGCACAGA reverse: 3'-AGAAGCAAGATGGAAGTCGAC, probe: 5'-
144 ACCAAATAACGCTGGGTC) of *MLH1* gene promoter were designed to specially amplify the CpG sites and
145 to be located up-stream of exon one (NG_007109.2 region: 4,591 -4,710), which was interpreted as the
146 promoter regions of the *MLH1* gene using ABI Primer Express version 3.0 (Thermofisher, Waltham MA).
147 The primers and probe were supplied by ABI and optimized for the present study. Briefly, qPCR using
148 FastSYBR Green Master Mix (cat # 4385610, ThermoFisher, Waltham MA) was used to check for

149 nonspecific product formation by dissociation-curve while qPCR using TaqMan Universal Master Mix II
150 (Cat # 4440038, ThermoFisher, Waltham, MA) was used to test the probes and determine the qPCR
151 efficiency.

152

153 **Real-Time PCR (qPCR)**

154 qPCR was performed in accordance with the Minimum Information for Publication of Quantitative Real-
155 Time PCR Experiments (MIQE) guidelines [13]. The reaction volume was set as 20 μ L and each reaction
156 contained 1 \times Taqman Universal PCR mastermix (cat # 4304437, Life Technologies, Carlsbad CA), 900nM
157 final concentration forward and reverse primers, 200nM Probes and 6ng template DNA. At least one
158 PCR No Template Control (NTC) was run for each assay as controls on all plates. All qPCR analyses were
159 performed in triplicate on a real time PCR system (Model 7500 Fast, Thermo Fisher Scientific®). qPCR
160 thermal cycling conditions were as follows: 95°C for 10min, followed by 40 cycles of 95°C for 15s, 60°C
161 for 1 min. The SDS software v2.4 (Life Technologies) was used to calculate the quantification cycle (Cq)
162 value. Invalid PCR results indicated by high cycle threshold (CT) values > 33 were omitted from the
163 analysis.

164

165 **Droplet digital PCR (ddPCR)**

166 ddPCR was performed in accordance with the Minimum Information for Publication of Digital
167 Quantitative PCR Experiments (digital MIQE) guidelines [14] using Bio-Rad QX200 droplet digital system
168 (Bio-Rad, Hercules, CA, USA). Droplets were generated in the QX200 automated droplet generator,
169 amplified on an Applied Biosystems VeritiDx thermal cycler using standard ddPCR cycling conditions with
170 an anneal/extension temperature of 55°C and read using the QX200 droplet reader (Bio-Rad). The
171 analysis was performed using QuantaSoft software (v1.6.6.0320, Bio-Rad, Pleasanton, CA) as follows.
172 Positive droplets, containing amplification products, were discriminated from negative droplets without

173 amplification products by applying a fluorescence amplitude threshold. The threshold was set manually,
174 using the 1D amplitude chart. To assess the specificity of the ddPCR assay, initial experiments were
175 conducted using 100% methylated DNA and 100% unmethylated DNA.

176

177 **Data analysis**

178 Values are expressed as mean \pm SD, unless otherwise stated. Differences between groups were tested using t
179 test, Mann-Whitney U test, or ANOVA, where appropriate.

180 Assay performance parameters (sensitivity, specificity and ROC curves) were calculated using the 28

181 cancer samples and 20 healthy donors. To determine the optimal DNA input amount of the qPCR assay, the

182 standard DNA sample was first prepared as 100 ng μL^{-1} and then was gradient diluted from 100ng μL^{-1} to 1ng

183 μL^{-1} . For the optimal input amount of the ddPCR assay, the DNA sample was gradient diluted from 6ng μL^{-1} to

184 0.094ng μL^{-1} . Each assay was conducted in at least three independent runs with four replicates in a run. The

185 limit of detection (LOD) is defined as the lowest detectable percentage of methylation above background and

186 was determined in the ddPCR assay by using the methylation spike-in samples. The linear regression analyses

187 were performed to obtain the correlation coefficients coefficient of determination (R^2) and the LOD. Receiver

188 operating characteristic (ROC) curves were calculated to establish the cutoffs or threshold methylation level

189 of *MLH1* promoter using the Web-based calculator for ROC curves from online tool

190 <http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html> which calculates the cutoff based on the value

191 corresponding with the highest average of sensitivity and specificity. Area under the curve (AUC) indicates the

192 value of the test. Correlation between variables was tested using Spearman's rank correlation. The Bonferroni

193 correction was used in case of multiple pairwise comparisons.

194

195 **Results**

196 **Assay performance**

197 FastSYBR Green PCR revealed a single peak of amplified product following melt curve analysis and
198 TaqMan qPCR efficiency of 94.98% (**Figure 1a**) with an input range of 100ng-3.125ng. The qPCR assay
199 showed very good linearity >0.99 indicating good performance of the assay across this range of input.
200 The linearity of the ddPCR assay was determined by titrating the input of sheared methylated control
201 DNA and sheared unmethylated control DNA from 6ng to 0.096ng and measuring the number of positive
202 droplets detected. In ddPCR assay, using sheared synthetic DNA, *MLH1* methylation was able to be
203 detected using as low as 0.096ng of DNA (**Figure 1b**).

204 To determine the optimal input amount for this assay, a titration of input amounts was tested
205 from 4ng DNA to 0.125ng DNA in two-fold dilutions. Four different conditions were tested including
206 methylated DNA treated with *HpaII* enzyme, methylated DNA untreated, unmethylated DNA treated
207 with *HpaII* enzyme, and unmethylated DNA untreated. With 0.125-4ng input of synthetic DNA, *MLH1*
208 methylation levels detected in the methylated treated, methylated untreated, and unmethylated
209 untreated sample remained fairly constant, the unmethylated enzyme treated sample showed a
210 significant reduction in signal ($P < 0.001$, **Figure 2**). To achieve the optimal signal/noise ratio and
211 operational/practical advantage, 1ng was selected as the optimal input for further analysis.

212 In terms of the percent methylation, the LOD of this assay was determined by spiking different
213 percentages of methylated DNA into normal plasma, extracting cfDNA and then testing 1ng of DNA in
214 ddPCR to detect *MLH1* methylation levels. While the entire titration showed a good linear relationship
215 overall ($R^2 = 0.985$, **Figure 3**), levels below 10% methylation displayed a weak linear relationship ($R^2 =$
216 0.427) when examined alone. Therefore, we can consider 10% methylation, or approximately 8 positive
217 droplets, as the LOD for the ddPCR assay.

218

219 **Testing *MLH1* methylation in healthy controls and CRC patient samples**

220 We tested 28 advanced stage CRC cancer samples and 20 normal healthy samples in ddPCR assay to
221 assess the level of *MLH1* methylation in each group (**Figure 4**). Except for one sample that failed due to
222 insufficient amount of cfDNA, the rest of samples returned quality ddPCR results. First, both groups
223 when untreated showed a significant increase in positive droplets over the treated samples ($P < 0.001$).
224 Normal control samples without enzyme treatment yielded 86.45 ± 41.45 positive droplets for *MLH1*
225 while cancer samples without enzyme treatment yielded an average of 117.39 ± 43.60 positive droplets
226 for *MLH1* ($P > 0.05$). Normal healthy samples treated with *HpaII* enzyme yielded 1.63 ± 1.84 positive
227 droplets indicating that the majority of *MLH1* in normal samples is unmethylated. In contrast the cancer
228 samples treated with *HpaII* enzyme yielded an average of 15.91 ± 9.76 positive droplets suggesting
229 higher methylation levels in the cancer population. A t-test of the cancer and healthy volunteers'
230 samples was shown to be a statistically significant difference ($P < 0.001$).
231 The data generated here could be used to create a receiver operating characteristic (ROC) curve to
232 differentiate cancer from normal patients with an area under the curve (AUC) of 0.965. The positive
233 droplet count for the normal and CRC patients could then be converted to percent methylation using
234 the equation of the line of best fit from the methylation spike-in standard curve (**Figure 5a**). In our
235 cohort all normal patient samples fell below the limit of detection and 21/27 (77.8%) of cancer patients
236 were above this threshold.

237 By correlating promoter methylation status with age, an increase of methylation levels of *MLH1*
238 with aging observed in CRC samples ($R = 0.21$, $P = 0.29$) (Figure 6a). Females showed higher *MLH1*
239 promoter methylation levels with respect to males ($P = 0.23$) (Fig. 6b).

240

241 **Cut-off value determination strategies**

242 Based on the results of the ddPCR assay, we set up the cut-off values of *MLH1* promoter methylation for
243 discriminating the healthy cases and the malignant cases. Data showed that the optimal sensitivity and

244 specificity was displayed by the assay when using 5 positive droplets as cutoff value (88% sensitivity and
245 95% specificity). However, this cut-off was below the reported LOD of the assay. Therefore, the LOD, 8
246 positive droplets was used as an optimal cut-off to distinguish healthy from CRC patients (78% sensitivity
247 and 100% specificity, **table 1**).

248 A tiered-based (20th, 50th, 80th percentile) cutoff was applied to distinguish CRC samples with
249 different methylation level by using the spike-in normal plasma data. A cut-off at the 80th percentile
250 could identify those patients with over 40% *MLH1* methylation, those samples falling in between the
251 50th percentile and 80th would be identified as 25-40% *MLH1* methylation and those samples below the
252 20th percentile could be considered as 'low' *MLH1* methylation (<10% methylation) (**Figure 5b**).

253 However, further clinical analysis of samples would be needed to elucidate the optimal percentiles of
254 the cut-off to ensure a meaningful analysis.

255

256 **Discussion**

257 Given the stability of DNA methylation in body fluids, cfDNA methylation markers using plasma samples
258 are a promising matrix for non-invasive clinical biomarker implementation [15]. In addition, global
259 hypomethylation of the human genome paired with CpG islands of hypermethylation, allow for a
260 minimal number of loci to be targeted and yet have sufficient test coverage [16]. These features make
261 cancer-associated DNA methylation changes an attractive source for clinical biomarker discovery[17,
262 18]. Therefore, targeting DNA methylation changes as cancer biomarkers hold potential for detection
263 and minimal residual disease. However, for such assays to be adopted into a clinical setting, high
264 sensitivity and specificity need to be demonstrated. In many cases it is crucial for the assay to detect
265 only a few cancer specific DNA methylation changes in a large background of normal DNA methylation

266 patterns. We developed a MSRE digested ddPCR assay to quantitate the methylation of *MLH1* promoter in
267 cfDNA for detection of low concentrations of methylated DNA fragments in advanced stage CRC.

268 While many other ctDNA methylation biomarkers for cancer have been reported [9, 19-21],
269 these mainly depend on the methylation level of individual CpG sites, which limits the detection
270 sensitivity. Until now, only one blood-based assay that detects *SEPT9* gene methylation was approved by
271 the U.S. Food and Drug Administration for CRC screening with a sensitivity of ~70% and a specificity of
272 ~80% or above [22]. Our assay tests the methylation level of the *MLH1* promoter region and offer a cost
273 efficient, clinical applicable alternative approach compared to bisulfite-treatment approaches. A MSRE
274 assay followed by ddPCR achieves absolute quantification of the degree of methylation of the DNA with
275 minimal (1ng) cfDNA input and the assay does not rely on a standard calibration curve. Using the control
276 DNA samples, we have proven that the LOD of the ddPCR assay is supersensitive, with 25-fold
277 improvement compared to the conventional qPCR-based assay. Based on the results of the ddPCR assay,
278 ROC curve produced from the data gave an AUC of 0.965 indicating the assay does a good job of
279 differentiating normal samples from cancer samples.

280 Moreover, we applied comprehensive approaches to determine the cut-off value for potential
281 clinical applications. First, a statistical cut-off of 8 positive droplets with optimal sensitivity and
282 specificity (78% sensitivity and 100% specificity) of the assay for discriminating healthy volunteers from
283 CRC patients was determined. Second, a tiered cut-off was applied to distinguish CRC samples with
284 different methylation levels (20%, 50%, 80%). Coppede et al. demonstrated that *MLH1* promoter level in
285 CRC tissue is highly correlated with MLH1 protein expression using immunostaining [23]. Their results
286 showed that in the vast majority of cases (80%) if promoter methylation was present, MLH1
287 immunostaining was negative, and, in contrast, in 93% of cases with 0% promoter methylation MLH1
288 immunostaining was positive. In the 20% of cases with both positive MLH1 promoter methylation and
289 positive immunostaining, the overall promoter methylation level was low [23]. Therefore, it is

290 reasonable to speculate that the tiered *MLH1* methylation level is linked to MLH1 protein expression
291 and could be further developed as a complementary assay of IHC tests for mismatch repair deficient
292 (dMMR) to identify the MSI-H or dMMR tumor status.

293 Our analysis revealed an increased ctDNA *MLH1* promoter methylation in plasma with aging, as
294 well as gender differences with females showing higher hMLH1 promoter methylation than males in CRC
295 patients (**Figure 6**). Although there is no statistical significance due to the small sample size, our results
296 showed the similar trend observed in CRC tissue as previous reported [23]. These findings provide
297 evidence that *MLH1* promoter methylation signals could be captured in plasma. Recent publications
298 have established an association between the epigenetic alterations found in primary tumor specimens
299 and in plasma, suggesting epigenetic biomarkers from liquid biopsy can be implemented as surrogate
300 tumor biomarkers [24-28].

301 The tissue based pooled frequency of *MLH1* promoter methylation in CRC is reported as ~18% in
302 sporadic CRC [29, 30]. DNA methylation of *MLH1* has been found in more than 80% of sporadic CRC with
303 MSI and predict the mutation load. Thus, the assessment of *MLH1* promoter methylation status will add
304 essential diagnostic and prognostic information with clinical relevance. In our study, a theoretical cut-off
305 methylation levels based on real-world data was set at ~24 positive ddPCR droplets (equivalent to 40%
306 methylation in **Figure 5b**) in CRC plasma samples. As the current study is relatively small, proper
307 validation using a different cohort with a larger sample size is crucial. Comprehensive and fit-for-
308 purpose cut-off determination strategies need to be considered to fulfill the different project goals and
309 clinical application specific requirements.

310

311

312

313 **Conclusion**

314 The developed liquid biopsy assay for detecting *MLH1* promoter methylation by ddPCR is a convenient
315 and cost-efficient approach which could be easily transferred to a clinical setting. The simple MSRE
316 ddPCR procedure and high sensitivity of the assay provides a reliable ctDNA methylation biomarker
317 assay to reflect single or multiple DNA methylation changes originating in tumor cells. Furthermore, this
318 approach can be applied to other genes of interest, or cancer indications, where methylation patterns
319 could reveal crucial clinically relevant information for future clinical biomarker and/or companion
320 diagnostic development and implementation.

321 **Declarations**

322

323 ***Ethics approval and consent to participate***

324 The written informed consent was received from all subjects used in this study and all samples were
325 collected under IRB approved protocols, as defined by the specimen provider BioIVT. Ethical approval
326 was obtained from BioIVT ethics committee.

327

328 ***Consent for publication***

329 Not applicable

330

331 ***Availability of data and materials***

332 The datasets used and/or analysed during the current study are available from the corresponding author
333 on reasonable request.

334

335 ***Competing interests***

336 The authors declare that they have no competing interests.

337

338 ***Funding***

339 This project was supported by the EMD Serono, Merck KGaA, Darmstadt, Germany

340

341 **Author Contributions**

342 DW contributed in the planning and execution of experiments, data analysis and writing of article.

343 DO conducted in house experiments, initial data analysis, and writing of article.

344 JSG contributed to data analysis.

345 TC provided scientific expertise in study planning and manuscript drafting.

346 JS provided scientific expertise.

347 ZF provided scientific expertise, advised in experimental planning and data analysis as well as substantial
348 contribution to composing the manuscript.

349

350 **Acknowledgements**

351 The authors would like to acknowledge Waldemar Miller and Xiaoyan Yin for their biostatistical analysis
352 and guidance.

353

354 **References**

- 355 1. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M,
356 Webster M *et al*: **The human colon cancer methylome shows similar hypo- and**
357 **hypermethylation at conserved tissue-specific CpG island shores.** *Nat Genet* 2009, **41**(2):178-
358 186.
- 359 2. Chen W, Swanson BJ, Frankel WL: **Molecular genetics of microsatellite-unstable colorectal**
360 **cancer for pathologists.** *Diagnostic pathology* 2017, **12**(1):24.
- 361 3. Boland CR, Goel A: **Microsatellite instability in colorectal cancer.** *Gastroenterology* 2010,
362 **138**(6):2073-2087.e2073.
- 363 4. Yamamoto H, Imai K: **Microsatellite instability: an update.** *Archives of toxicology* 2015,
364 **89**(6):899-921.
- 365 5. Liu L, Yang X: **Implication of Reprimo and hMLH1 gene methylation in early diagnosis of gastric**
366 **carcinoma.** *International journal of clinical and experimental pathology* 2015, **8**(11):14977-
367 14982.
- 368 6. Cabel L, Proudhon C, Romano E, Girard N, Lantz O, Stern MH, Pierga JY, Bidard FC: **Clinical**
369 **potential of circulating tumour DNA in patients receiving anticancer immunotherapy.** *Nature*
370 *reviews Clinical oncology* 2018, **15**(10):639-650.
- 371 7. Hu G, Qin L, Zhang X, Ye G, Huang T: **Epigenetic Silencing of the MLH1 Promoter in Relation to**
372 **the Development of Gastric Cancer and its use as a Biomarker for Patients with Microsatellite**
373 **Instability: a Systematic Analysis.** *Cellular physiology and biochemistry : international journal of*
374 *experimental cellular physiology, biochemistry, and pharmacology* 2018, **45**(1):148-162.
- 375 8. Gai W, Sun K: **Epigenetic Biomarkers in Cell-Free DNA and Applications in Liquid Biopsy.** *Genes*
376 2019, **10**(1).

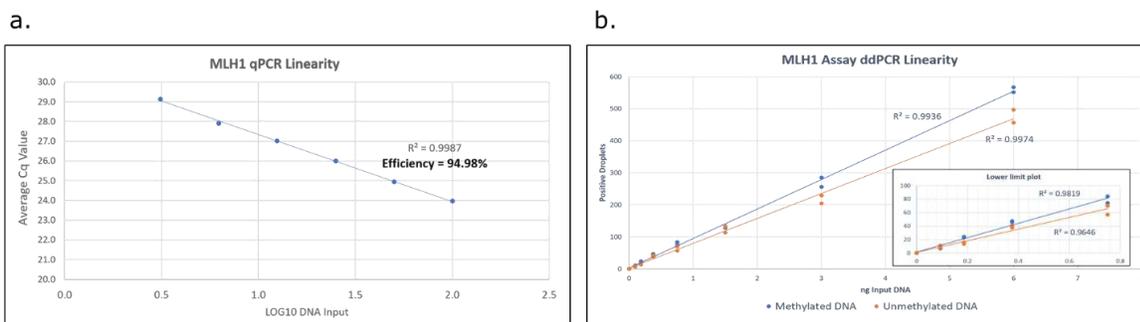
- 377 9. Warton K, Mahon KL, Samimi G: **Methylated circulating tumor DNA in blood: power in cancer**
378 **prognosis and response.** *Endocr Relat Cancer* 2016, **23**(3):R157-171.
- 379 10. Lissa D, Robles AI: **Methylation analyses in liquid biopsy.** *Translational lung cancer research*
380 2016, **5**(5):492-504.
- 381 11. Willis J, Lefterova MI, Artyomenko A, Kasi PM, Nakamura Y, Mody K, Catenacci DVT, Fakhri M,
382 Barbacioru C, Zhao J *et al*: **Validation of Microsatellite Instability Detection Using a**
383 **Comprehensive Plasma-Based Genotyping Panel.** *Clin Cancer Res* 2019, **25**(23):7035-7045.
- 384 12. Liu X, Ren J, Luo N, Guo H, Zheng Y, Li J, Tang F, Wen L, Peng J: **Comprehensive DNA methylation**
385 **analysis of tissue of origin of plasma cell-free DNA by methylated CpG tandem amplification**
386 **and sequencing (MCTA-Seq).** *Clin Epigenetics* 2019, **11**(1):93.
- 387 13. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW,
388 Shipley GL *et al*: **The MIQE guidelines: minimum information for publication of quantitative**
389 **real-time PCR experiments.** *Clin Chem* 2009, **55**(4):611-622.
- 390 14. Huggett JF, Foy CA, Benes V, Emslie K, Garson JA, Haynes R, Hellemans J, Kubista M, Mueller RD,
391 Nolan T *et al*: **The digital MIQE guidelines: Minimum Information for Publication of**
392 **Quantitative Digital PCR Experiments.** *Clin Chem* 2013, **59**(6):892-902.
- 393 15. Galanopoulos M, Tsoukalas N, Papanikolaou IS, Tolia M, Gazouli M, Mantzaris GJ: **Abnormal**
394 **DNA methylation as a cell-free circulating DNA biomarker for colorectal cancer detection: A**
395 **review of literature.** *World J Gastrointest Oncol* 2017, **9**(4):142-152.
- 396 16. Pfeifer GP: **Defining Driver DNA Methylation Changes in Human Cancer.** *International journal of*
397 *molecular sciences* 2018, **19**(4).
- 398 17. Locke WJ, Guanzon D, Ma C, Liew YJ, Duesing KR, Fung KYC, Ross JP: **DNA Methylation Cancer**
399 **Biomarkers: Translation to the Clinic.** *Front Genet* 2019, **10**:1150.
- 400 18. Taryma-Leśniak O, Sokolowska KE, Wojdacz TK: **Current status of development of methylation**
401 **biomarkers for in vitro diagnostic IVD applications.** *Clin Epigenetics* 2020, **12**(1):100.
- 402 19. Leygo C, Williams M, Jin HC, Chan MWY, Chu WK, Grusch M, Cheng YY: **DNA Methylation as a**
403 **Noninvasive Epigenetic Biomarker for the Detection of Cancer.** *Disease markers* 2017,
404 **2017**:3726595.
- 405 20. Dong L, Ren H: **Blood-based DNA Methylation Biomarkers for Early Detection of Colorectal**
406 **Cancer.** *Journal of proteomics & bioinformatics* 2018, **11**(6):120-126.
- 407 21. Nunes SP, Moreira-Barbosa C, Salta S, Palma de Sousa S, Pousa I, Oliveira J, Soares M, Rego L,
408 Dias T, Rodrigues J *et al*: **Cell-Free DNA Methylation of Selected Genes Allows for Early**
409 **Detection of the Major Cancers in Women.** *Cancers (Basel)* 2018, **10**(10).
- 410 22. Semaan A, van Ellen A, Meller S, Bergheim D, Branchi V, Lingohr P, Goltz D, Kalff JC, Kristiansen
411 G, Matthaei H *et al*: **SEPT9 and SHOX2 DNA methylation status and its utility in the diagnosis of**
412 **colonic adenomas and colorectal adenocarcinomas.** *Clin Epigenetics* 2016, **8**:100.
- 413 23. Coppedè F, Migheli F, Lopomo A, Failli A, Legitimo A, Consolini R, Fontanini G, Sensi E, Servadio
414 A, Seccia M *et al*: **Gene promoter methylation in colorectal cancer and healthy adjacent**
415 **mucosa specimens: correlation with physiological and pathological characteristics, and with**
416 **biomarkers of one-carbon metabolism.** *Epigenetics* 2014, **9**(4):621-633.
- 417 24. Li W, Zhang X, Lu X, You L, Song Y, Luo Z, Zhang J, Nie J, Zheng W, Xu D *et al*: **5-**
418 **Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for**
419 **human cancers.** *Cell research* 2017, **27**(10):1243-1257.
- 420 25. Chan KC, Jiang P, Chan CW, Sun K, Wong J, Hui EP, Chan SL, Chan WC, Hui DS, Ng SS *et al*:
421 **Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number**
422 **aberrations by plasma DNA bisulfite sequencing.** *Proc Natl Acad Sci U S A* 2013, **110**(47):18761-
423 18768.

- 424 26. Guo S, Diep D, Plongthongkum N, Fung HL, Zhang K, Zhang K: **Identification of methylation**
425 **haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-**
426 **origin mapping from plasma DNA.** *Nat Genet* 2017, **49**(4):635-642.
- 427 27. Schwarzenbach H, Hoon DS, Pantel K: **Cell-free nucleic acids as biomarkers in cancer patients.**
428 *Nat Rev Cancer* 2011, **11**(6):426-437.
- 429 28. Krishnamurthy N, Spencer E, Torkamani A, Nicholson L: **Liquid Biopsies for Cancer: Coming to a**
430 **Patient near You.** *J Clin Med* 2017, **6**(1).
- 431 29. Li X, Yao X, Wang Y, Hu F, Wang F, Jiang L, Liu Y, Wang D, Sun G, Zhao Y: **MLH1 promoter**
432 **methylation frequency in colorectal cancer patients and related clinicopathological and**
433 **molecular features.** *PLoS One* 2013, **8**(3):e59064.
- 434 30. Shen S CX, Li H, Sun L, Yuan Y: **MLH1 Promoter Methylation and Prediction/Prognosis of Gastric**
435 **Cancer: A Systematic Review and Meta and Bioinformatic Analysis.** *J Cancer* 2018,
436 **9**((11)):1932–1942.
- 437

438 **Table 1.** Summary of sensitivity and specificity calculations at different positive droplets cut off.

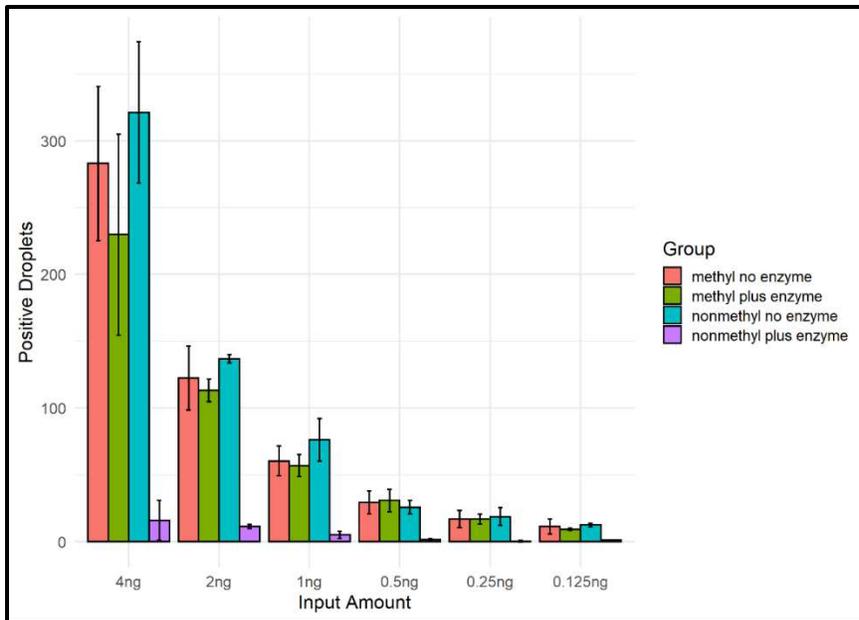
CUTOFF	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Sensitivity	0.96	0.96	0.93	0.93	0.93*	0.93	0.89	0.78#	0.78	0.70	0.67	0.67	0.52	0.48	0.44	0.44	0.44	0.41	0.26	0.26
Specificity	0.50	0.60	0.85	0.90	0.95*	0.95	0.95	1.00#	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ACCURACY	0.77	0.81	0.89	0.91	0.94	0.94	0.91	0.87	0.87	0.83	0.81	0.81	0.72	0.70	0.68	0.68	0.68	0.66	0.57	0.57
PPV	0.72	0.76	0.89	0.93	0.96	0.96	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

439
 440 * The optimal sensitivity and specificity were displayed at 5 positive droplets as cutoff value (88%
 441 sensitivity and 95% specificity).
 442 # As 5 positive droplets was below the reported LOD, 8 positive droplets were used as an exploratory cut-
 443 off to distinguish healthy from CRC patients (78% sensitivity and 100% specificity).
 444



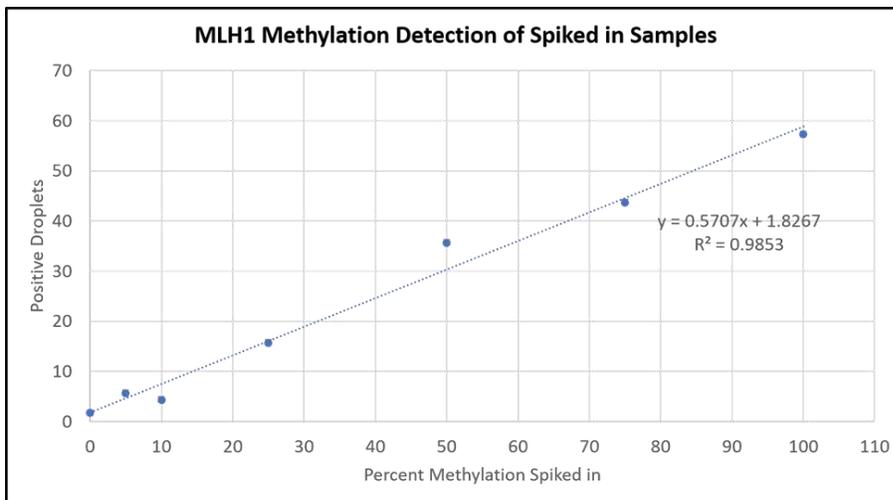
445
 446 **Figure 1. Titration of input amount using synthetic DNA to determine minimum input for qPCR and**
 447 **ddPCR. a.** Linearity of MLH1 qPCR assay showed very good linearity 0.9987 and efficiency 94.98% with
 448 an input range of 100ng-3.125ng control DNA. **b.** Comparative analysis of MLH1 promotor methylation
 449 using ddPCR demonstrated the ability to detect MLH1 copies in this assay at very low input level (<1ng
 450 equivalent input). As low as 0.096ng of MLH1 methylation DNA could be detected by ddPCR.

451
 452



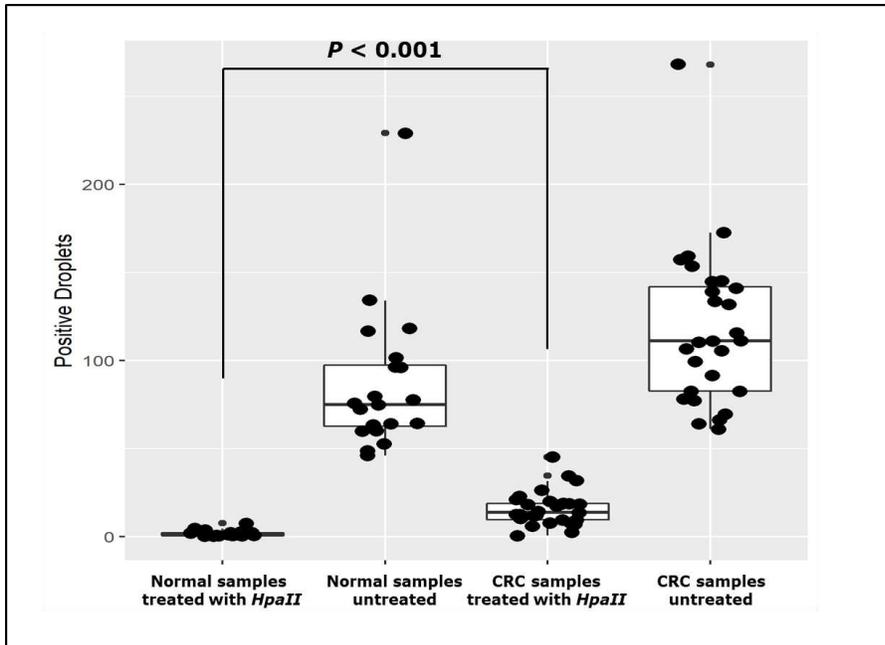
453
 454 **Figure 2. Sample input optimization for *MLH1* promoter methylation ddPCR assay.** With 0.125-4ng
 455 input of synthetic DNA, *MLH1* methylation levels detected in the methylated treated, methylated
 456 untreated, and unmethylated untreated sample remained fairly constant, the unmethylated enzyme
 457 treated sample showed a significant reduction in signal ($P < 0.001$). To achieve the optimal signal/noise
 458 ratio and operational/practical advantage, 1ng was selected as the optimal input for further analysis.

459



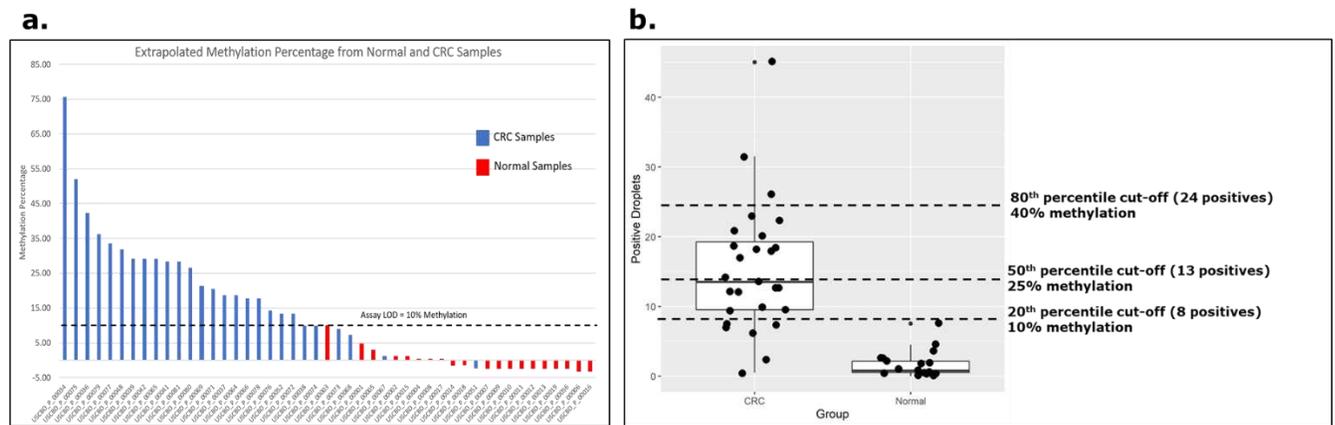
460
 461 **Figure 3. *MLH1* methylation detection of spike in normal plasma sample.** Standard curve generated by
 462 spiking in 50ng of DNA with known percentage of methylation into normal plasma, extracting cfDNA
 463 from the sample and testing in ddPCR. The normal samples had a measured baseline level of 0 positive
 464 droplets. The standard curve gave a R^2 value of 0.985 and the equation of best can be used to relate
 465 positive droplet count to percent methylation.

466



467

468 **Figure 4: Distribution of detected positive droplets for *MLH1* methylation in Normal and CRC Samples.**
 469 With 1ng input cfDNA, 20 normal and 28 CRC samples were treated or untreated with *HpaII* enzyme. In
 470 both populations the untreated samples showed higher positive droplets than the treated samples.
 471 Compared to normal samples, significant higher methylation levels in the cancer population ($P < 0.001$).
 472

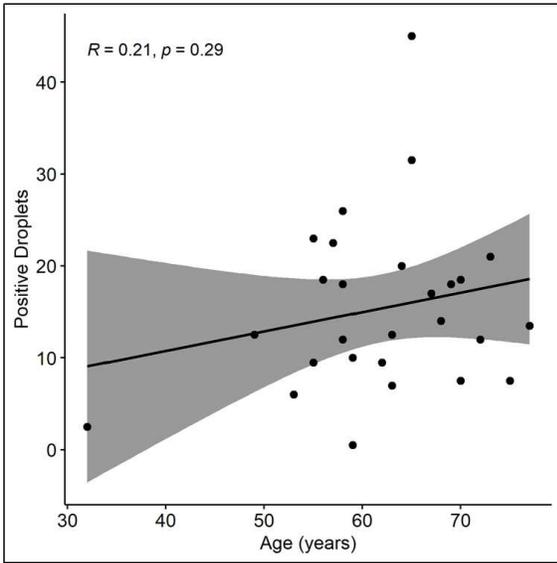


473

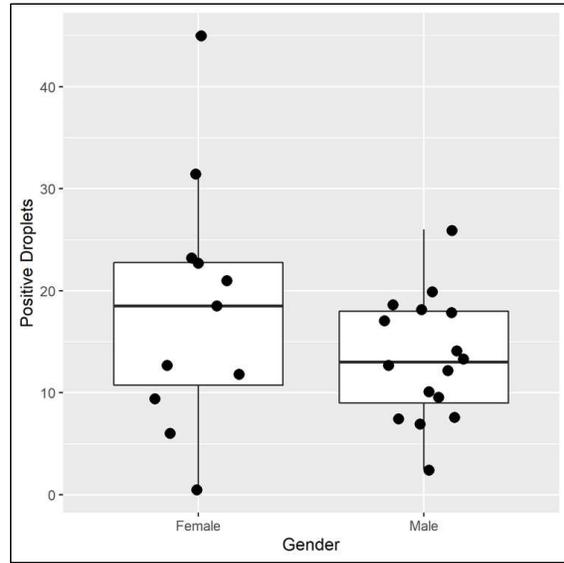
474 **Figure 5: Determining an optimal cut-off to distinguish healthy from CRC patients. a.** Based on ddPCR
 475 LOD at 10% methylation (approximately 8 positive droplets), all 20 normal patient samples fell below
 476 the limit of detection and 21/27 (77.8%) of cancer patients were above this threshold. **b.** A tiered-based
 477 cutoff (20th, 50th, 80th percentile based) to distinguish “methylation high” CRC samples from
 478 “methylation low” CRC samples.

479

a.



b.



480

481

482

Figure 6: Assessment of the contribution of age (a) and gender (b) with respect to *MLH1* promoter methylation in CRC plasma samples.

Figures

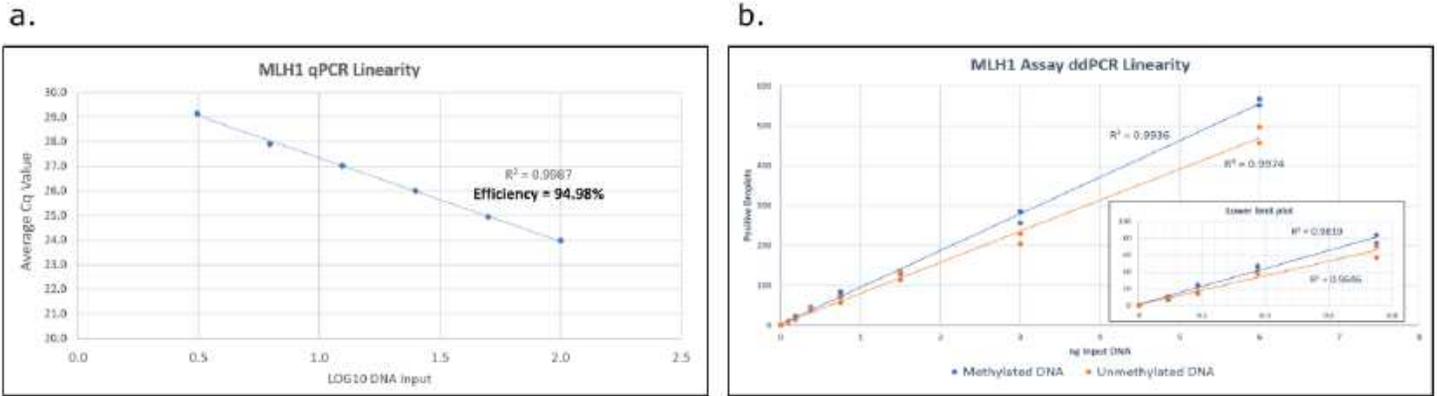


Figure 1

Titration of input amount using synthetic DNA to determine minimum input for qPCR and ddPCR. a. Linearity of MLH1 qPCR assay showed very good linearity 0.9987 and efficiency 94.98% within input range of 100ng-3.125ng control DNA. b. Comparative analysis of MLH1 promotor methylation using ddPCR demonstrated the ability to detect MLH1 copies in this assay at very low input level (<1ng equivalent input). As low as 0.096ng of MLH1 methylation DNA could be detected by ddPCR.

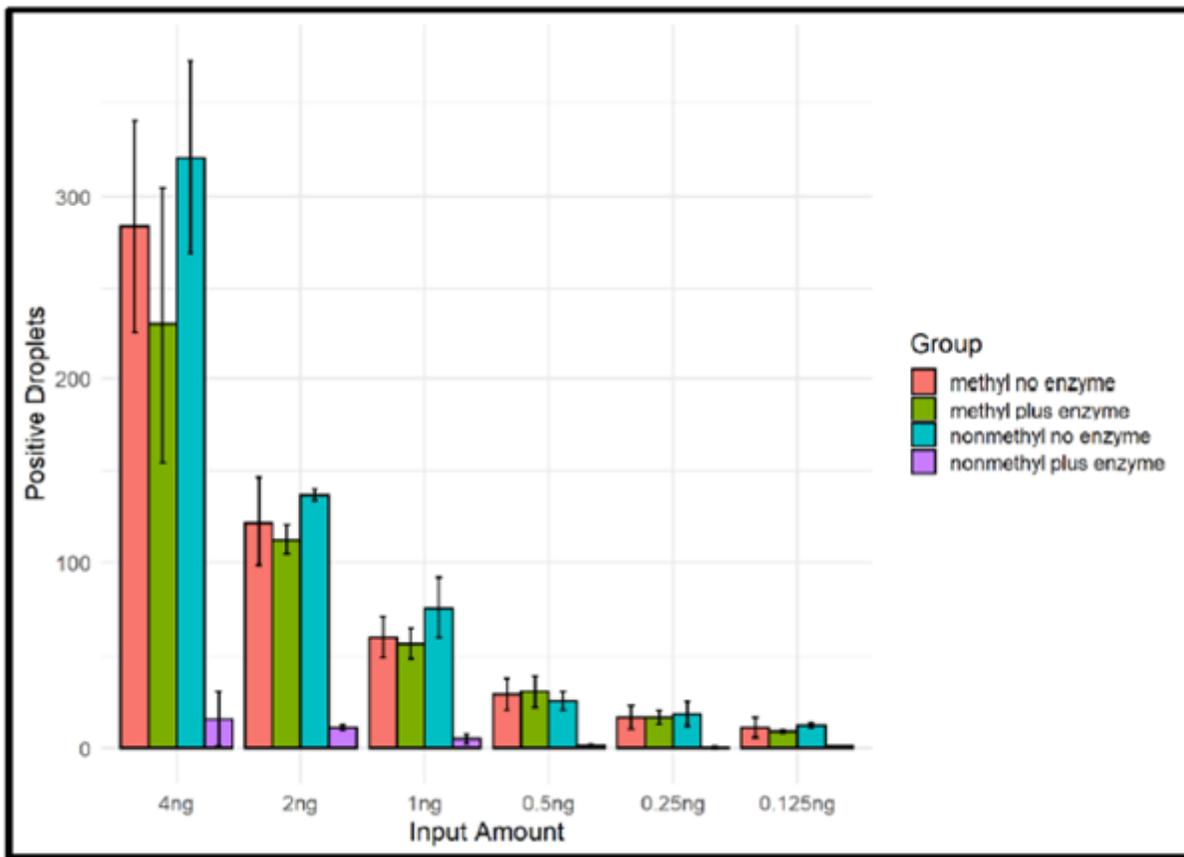


Figure 2

Sample input optimization for MLH1 promoter methylation ddPCR assay. With 0.125-4ng input of synthetic DNA, MLH1 methylation levels detected in the methylated treated, methylated untreated, and unmethylated untreated sample remained fairly constant, the unmethylated enzyme treated sample showed a significant reduction in signal ($P < 0.001$). To achieve the optimal signal/noise ratio and operational/practical advantage, 1ng was selected as the optimal input for further analysis.

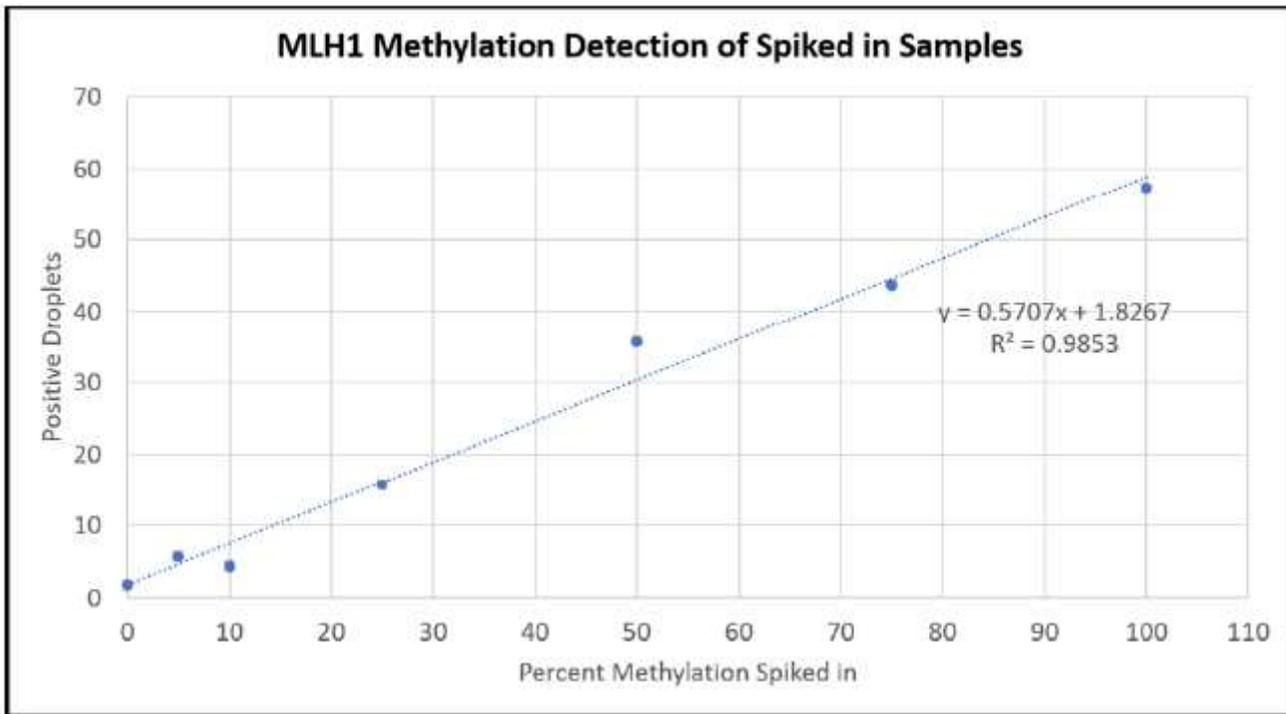


Figure 3

MLH1 methylation detection of spike in normal plasma sample. Standard curve generated by spiking in 50ng of DNA with known percentage of methylation into normal plasma, extracting cfDNA from the sample and testing in ddPCR. The normal samples had a measured baseline level of 0 positive droplets. The standard curve gave a R^2 value of 0.985 and the equation of best can be used to relate positive droplet count to percent methylation.

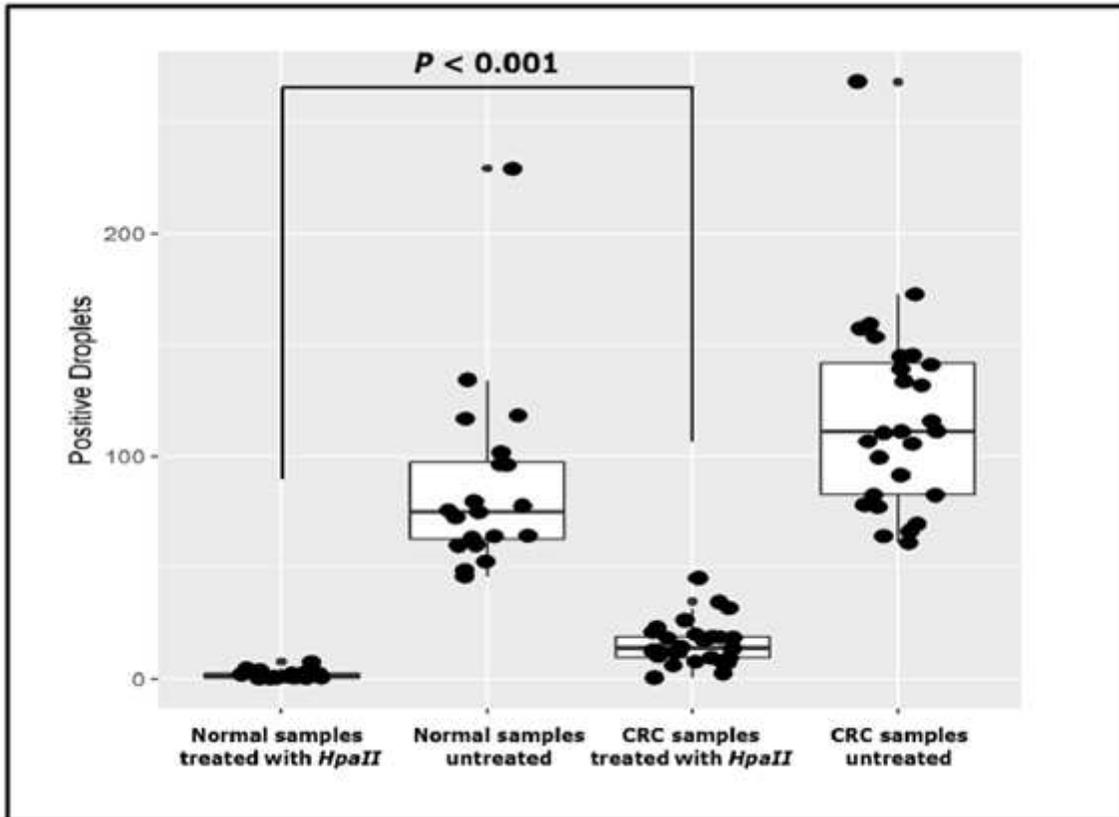


Figure 4

Distribution of detected positive droplets for MLH1 methylation in Normal and CRC Samples. With 1ng input cfDNA, 20 normal and 28 CRC samples were treated or untreated with HpaII enzyme. In both populations the untreated samples showed higher positive droplets than the treated samples. Compared to normal samples, significant higher methylation levels in the cancer population ($P < 0.001$).

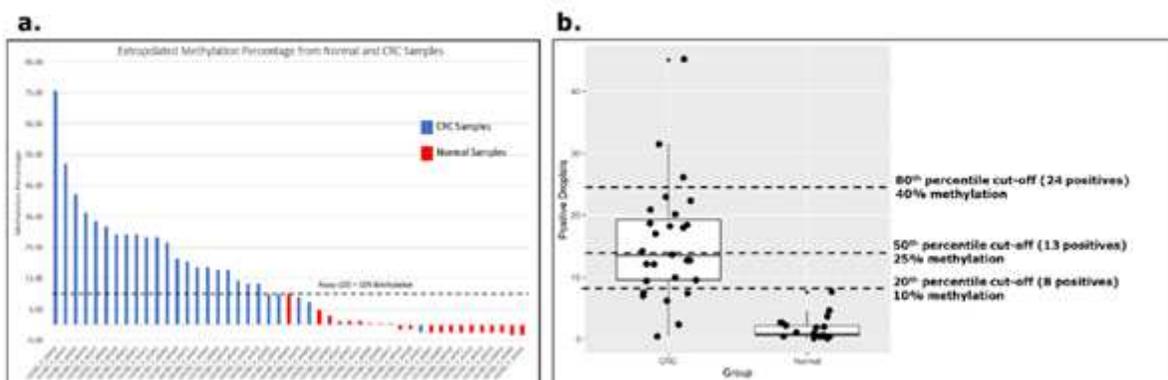


Figure 5

Determining an optimal cut-off to distinguish healthy from CRC patients. a. Based on ddPCR LOD at 10% methylation (approximately 8 positive droplets), all 20 normal patient samples fell below the limit of detection and 21/27 (77.8%) of cancer patients were above this threshold. b. A tiered-based cutoff (20th,

50th, 80th percentile based) to distinguish “methylation high” CRC samples from “methylation low” CRC samples.

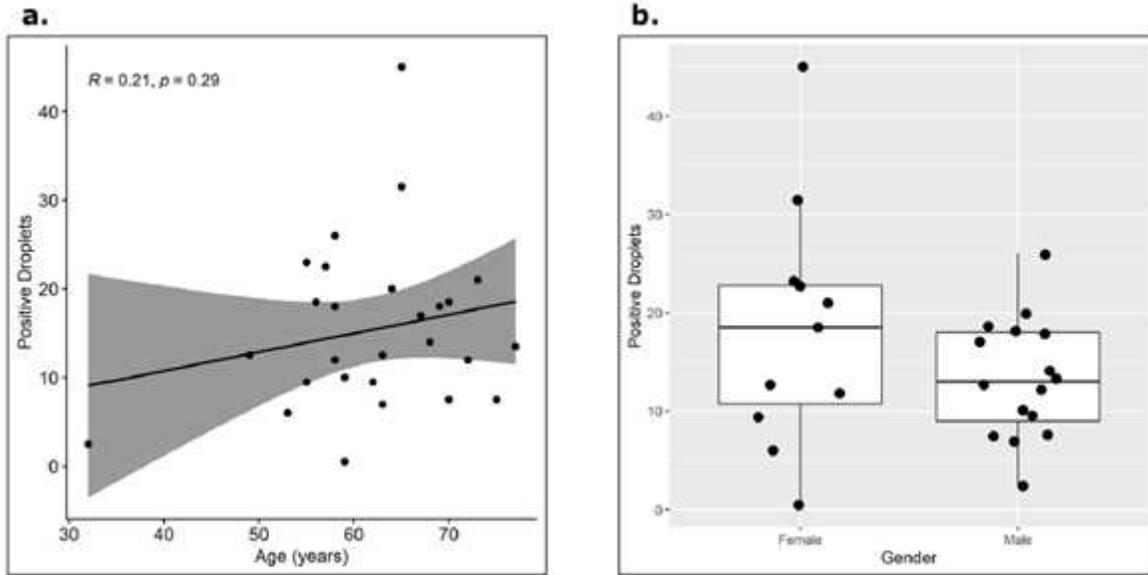


Figure 6

Assessment of the contribution of age (a) and gender (b) with respect to MLH1 promoter methylation in CRC plasma samples.