

# CpG island methylator phenotype-low (CIMP-low) colorectal cancer shows not only few methylated CIMP-high-specific CpG islands, but also low-level methylation at individual loci

Takako Kawasaki<sup>1</sup>, Mutsuko Ohnishi<sup>1</sup>, Katsuhiko Nosho<sup>1</sup>, Yuko Suemoto<sup>1</sup>, Gregory J Kirkner<sup>2</sup>, Jeffrey A Meyerhardt<sup>1,2</sup>, Charles S Fuchs<sup>1,2</sup> and Shuji Ogino<sup>1,3</sup>

<sup>1</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA; <sup>2</sup>Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA and <sup>3</sup>Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

**The CpG island methylator phenotype (CIMP or CIMP-high) with widespread promoter methylation is a distinct phenotype in colorectal cancer. However, the concept of CIMP-low with less extensive CpG island methylation is still evolving. Our aim is to examine whether density of methylation in individual CpG islands was different between CIMP-low and CIMP-high tumors. Utilizing MethyLight technology and 889 population-based colorectal cancers, we quantified DNA methylation (methylation index, percentage of methylated reference) at 14 CpG islands, including 8 CIMP-high-specific loci (*CACNA1G*, *CDKN2A* (*p16*), *CRABP1*, *IGF2*, *MLH1*, *NEUROG1*, *RUNX3* and *SOCS1*). Methylation positivity in each locus was defined as methylation index > 4. Low-level methylation (methylation index > 0, < 20) in each CIMP-high-specific locus was significantly more common in 340 CIMP-low tumors (1/8–5/8 methylation-positive loci) than 133 CIMP-high tumors (≥ 6/8 methylation-positive loci) and 416 CIMP-0 tumors (0/8 methylation-positive loci) ( $P \leq 0.002$ ). In the other six loci (*CHFR*, *HIC1*, *IGFBP3*, *MGMT*, *MINT31* and *WRN*), which were not highly specific for CIMP-high, low-level methylation, was not persistently more prevalent in CIMP-low tumors. In conclusion, compared to CIMP-high and CIMP-0 tumors, CIMP-low colorectal cancers show not only few methylated CIMP-high-specific CpG islands, but also more frequent low-level methylation at individual loci. Our data may provide supporting evidence for a difference in pathogenesis of DNA methylation between CIMP-low and CIMP-high tumors.**

*Modern Pathology* (2008) 21, 245–255; doi:10.1038/modpathol.3800982; published online 18 January 2008

**Keywords:** colon cancer; CpG island methylator phenotype; DNA methylation; promoter; MethyLight; CIMP-low

Epigenetic aberrations are important mechanisms in human carcinogenesis.<sup>1</sup> A number of tumor suppressor genes are silenced by promoter methylation in colorectal cancers.<sup>1,2</sup> A subset of colorectal cancers have been shown to exhibit widespread promoter methylation, which is referred to as the CpG island methylator phenotype (CIMP).<sup>1,3</sup> CIMP-high colorectal tumors have a distinct clinical, pathologic and molecular profile, such as associations with proximal tumor location, female sex, poor

tumor differentiation, inactive WNT/ $\beta$ -catenin (CTNNB1) and high *BRAF* and low *TP53* mutation rates,<sup>4–6</sup> independent of microsatellite instability status.<sup>7–10</sup> Although controversial, CIMP may have prognostic implications in colorectal cancer.<sup>11–14</sup>

In contrast to CIMP-high in colorectal cancer, the concept of CIMP-low (with less widespread CIMP-specific promoter methylation) is still emerging.<sup>15,16</sup> While CIMP-high colorectal cancer is associated with female sex and *BRAF* mutation, CIMP-low is associated with *KRAS* mutation.<sup>17</sup> Thus, CIMP-low cannot be explained by a simple mixture of CIMP-high and CIMP-0 (CIMP-negative).<sup>17</sup> We have also demonstrated a possible link between CIMP-low, microsatellite instability-low, *MGMT* methylation and *KRAS* mutation in colorectal cancer.<sup>18</sup> 18q loss of heterozygosity is associated positively with

Correspondence: Dr S Ogino, MD, PhD, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA.

E-mail: shuji\_ogino@dfci.harvard.edu

Received 7 August 2007; revised 29 October 2007; accepted 31 October 2007; published online 18 January 2008

CIMP-0 and inversely with CIMP-low.<sup>19</sup> These data collectively support molecular differences between CIMP-low, CIMP-high and CIMP-0 in colorectal cancer. In light of molecular differences between CIMP-low and CIMP-high, we hypothesized that density of methylation in individual CpG islands might be different between CIMP-low and CIMP-high tumors. Examining DNA methylation levels in individual CpG islands may provide additional evidence for underlying molecular differences between CIMP-low and CIMP-high tumors.

In this study, using quantitative methylation analysis (MethyLight technology<sup>20</sup>) and a large number of population-based colorectal cancer samples, we have examined the prevalences of low-level methylation in 14 CpG islands among CIMP-low tumors, in comparison to CIMP-high and CIMP-0 tumors. MethyLight assays (quantitative methylation-specific PCR) are robust and reproducible in quantifying methylation in DNA from paraffin-embedded tumor tissue.<sup>21–27</sup> Low index of methylation in MethyLight analysis is not associated with gene silencing,<sup>24</sup> indicating that low methylation index in MethyLight indeed reflects low-level DNA methylation.

## Materials and methods

### Study Group

We utilized the databases of two large prospective cohort studies; the Nurses' Health Study ( $N=121\,700$  women followed since 1976),<sup>28</sup> and the Health Professional Follow-up Study ( $N=51\,500$  men followed since 1986).<sup>29</sup> Informed consent was obtained from all participants prior to inclusion in the cohorts. A subset of the cohort participants developed colorectal cancers during prospective follow-up. Thus, these colorectal cancers represented population-based, relatively unbiased samples (compared to retrospective or single-hospital-based samples). Previous studies on Nurses' Health Study and Health Professionals Follow-up Study have described baseline characteristics of cohort participants and incident colorectal cancer cases, and confirmed that our colorectal cancer cases were well representative as a population-based sample.<sup>28,29</sup> We collected paraffin-embedded tissue blocks from hospitals where cohort participants had undergone resections of primary colorectal cancers. We obtained specimens (hematoxylin and eosin (H&E)-stained slides, unstained slides and/or paraffin blocks) from 648 cases (retrieval rate 76%) in the Health Professionals Follow-up Study and 662 cases (retrieval rate 58%) in the Nurses' Health Study. We excluded cases if adequate paraffin-embedded tumor tissue with more than 40% tumor cellularity was not available, or if tumors were preoperatively treated by chemotherapy or radiation. On the basis of the availability of adequate tissue specimens, a total of 889 colorectal cancers (402 from the men's

cohort and 487 from the women's cohort) were included, while 421 tumors (246 men and 175 women) were excluded because of lack of adequate materials. Baseline characteristics among cohort participants with colorectal cancer who had or did not have available tumor tissue specimens were very similar.<sup>30</sup> Our tumor samples have been previously characterized for status of CIMP and methylation in various CpG islands.<sup>10,17,18</sup> However, we have not examined *CHFR* (checkpoint with forkhead and ring finger domains), *HIC1* (hypermethylated in cancer 1), *IGFBP3* (insulin-like growth factor binding protein 3), *MINT31* (methylated in tumor 31) and *WRN* (Werner syndrome). Moreover, in our previous studies,<sup>10,17,18</sup> we utilized percentage of methylated reference (methylation index) values only as the binary discriminator of methylation status (positive vs negative). Our current novel analysis of raw methylation index (percentage of methylated reference) values has given us new evidence for molecular differences between CIMP-low and CIMP-high as shown below. Tissue collection and analyses were approved by the Dana-Farber Cancer Institute and Brigham and Women's Hospital Institutional Review Boards.

### Histologic Evaluation of Tumor Cellularity

Tumor cellularity was visually estimated by a pathologist (SO) who examined H&E-stained slides under a light microscope. We excluded cases if adequate paraffin-embedded tumor tissue with 40% or more tumor cellularity was not available, or if tumors were previously treated by chemotherapy or radiation. Thus, tumor cellularity by itself should not have caused low-level of methylation in this study. We compared average tumor cellularities between CIMP-high, CIMP-low and CIMP-0 tumors, and confirmed there were no significant differences between these tumors (see the Results section).

### Sequencing of *KRAS* and *BRAF*

Genomic DNA extraction and whole genome amplification were performed as described.<sup>31</sup> PCR and Pyrosequencing targeted for *KRAS* codons 12 and 13, and *BRAF* codon 600 were performed using the PSQ96 HS System (Biotage, Uppsala, Sweden).<sup>17,31</sup>

### Bisulfite Conversion and Real-Time PCR (MethyLight) for Quantitative DNA Methylation Analysis

Sodium bisulfite treatment on tumor tissue samples and genomic DNA was performed as described previously.<sup>24</sup> We did not use laser capture microdissection because ample amounts of DNA were critical for reliable quantification of DNA methylation.<sup>24</sup> Utilizing MethyLight technology (real-time PCR)<sup>20</sup> and ABI 7300 (Applied Biosystems, Foster

City, CA, USA), we quantified methylation in 8 promoters (*CACNA1G* (calcium channel, voltage-dependent, T-type  $\alpha$ -1G subunit), *CDKN2A* (cyclin-dependent kinase inhibitor 2A (p16/INK4A)), *CRABP1* (cellular retinoic acid binding protein 1), *IGF2* (insulin-like growth factor 2), *MLH1*, *NEUROG1* (neurogenin 1), *RUNX3* (runt-related transcription factor 3) and *SOCS1* (suppressor of cytokine signaling 1)), all of which have been selected based on screening of 195 CpG islands.<sup>6,8</sup> We have previously shown that these eight markers are sensitive and specific for the detection of CIMP-high.<sup>10</sup> *COL2A1* (the collagen 2A1 gene) was used to normalize for the amount of input bisulfite-converted DNA.<sup>24</sup> We also examined six other CpG islands (*CHFR*, *HIC1*, *IGFBP3*, *MGMT*, *MINT31* and *WRN*) which were not in the eight-marker CIMP panel. PCR primers and probes for *CHFR*, *HIC1* and *MINT31* were (bisulfite-converted nucleotides are highlighted by bold face and italics): *CHFR*-F, 5'-CGG GAG**TTTT**TATGGGCGT-3' (GenBank no. AC127070, nucleotide nos. 62442–62460); *CHFR*-R, 5'-AACCGT CCCC**AAA**ACTACGAC-3' (nucleotide nos. 62525–62545); *CHFR*-probe, 6FAM-5'-CCTCGA**ACCG**CTCCA TCGAAATTCA-3'-BHQ-1 (nucleotide nos. 62482–62505); *HIC1*-F, 5'-**TCG**TTACGGTAGTCGTTGTTT**TC**-3' (GenBank no. L41919, nucleotide nos. 43–66); *HIC1*-R, 5'-**GAAA**ACTATCAACCCTCGATCGA-3' (nucleotide nos. 94–116); *HIC1*-probe, 6FAM-5'-**TCG**CGCGG**TCG**TCGTT**CG**-3'-BHQ-1 (nucleotide nos. 72–89); *MINT31*-F, 5'-**GTCG**TCGGCGT**TAT**TTTAGAAAGTT-3' (GenBank no. AC021491, nucleotide nos. 50060–50130); *MINT31*-R, 5'-CACCGACGCC**CAAC**ACA-3' (nucleotide nos. 50113–50130); *MINT31*-probe, 6FAM-5'-**ACG**CTCCGCTCCCGA**ATAC**CCA-3'-BHQ-1 (nucleotide nos. 50087–50107). Other primers and probes were described previously<sup>6</sup> (*IGFBP3* and *WRN*<sup>3</sup>). PCR conditions were 95°C for 10 min, followed by 45 cycles of 95°C for 15 s and 60°C for 1 min. A standard curve was made for each PCR plate by duplicated PCR amplifications for *COL2A1* on bisulfite-converted human genomic DNA at four different concentrations (in a five-fold dilution series).

The percentage of methylated reference (referred to as 'methylation index') at a specific locus was calculated by dividing the *GENE:COL2A1* ratio of template amounts in a sample by the *GENE:COL2A1* ratio of template amounts in *SssI*-treated human genomic DNA (presumably fully methylated) and multiplying this value by 100 (or  $100 \times 2/3$  for *CRABP1*, *IGF2*, *HIC1*, *IGFBP3*, *MINT31* and *WRN*).<sup>20</sup> A methylation index cutoff of 4 (methylation positive vs negative) was based on previously validated data.<sup>20,24</sup> There was no tumor with methylation index = 4, 10 or 20, exactly, at any locus. Precision and performance characteristics of bisulfite conversion and subsequent MethyLight assays have been previously evaluated and the assays have been validated.<sup>24</sup>

CIMP-high was defined as the presence of  $\geq 6$  methylated markers among the eight CIMP panel

promoters (*CACNA1G*, *CDKN2A*, *CRABP1*, *IGF2*, *MLH1*, *NEUROG1*, *RUNX3* and *SOCS1*), CIMP-low as the presence of 1/8 to 5/8 methylated markers, and CIMP-0 as 0/8 methylated markers, according to the previously validated criteria.<sup>10</sup>

### Statistical Analysis

For categorical data, we performed the  $\chi^2$ -test (or the Fisher's exact test when the number in any category is less than 10), using SAS program (Version 9.1, SAS Institute, Cary, NC, USA). The Wilcoxon's rank-sum test was used to compare the average frequencies of low-level methylation in all eight CpG islands between two tumor subtypes (CIMP-0 vs CIMP-low; or CIMP-low vs CIMP-high). The two sample *t*-test assuming unequal variances was used to compare average tumor cellularities between CIMP-high, CIMP-low and CIMP-0 tumors. All *P*-values were two-sided, and statistical significance was set at  $P \leq 0.05$ .

## Results

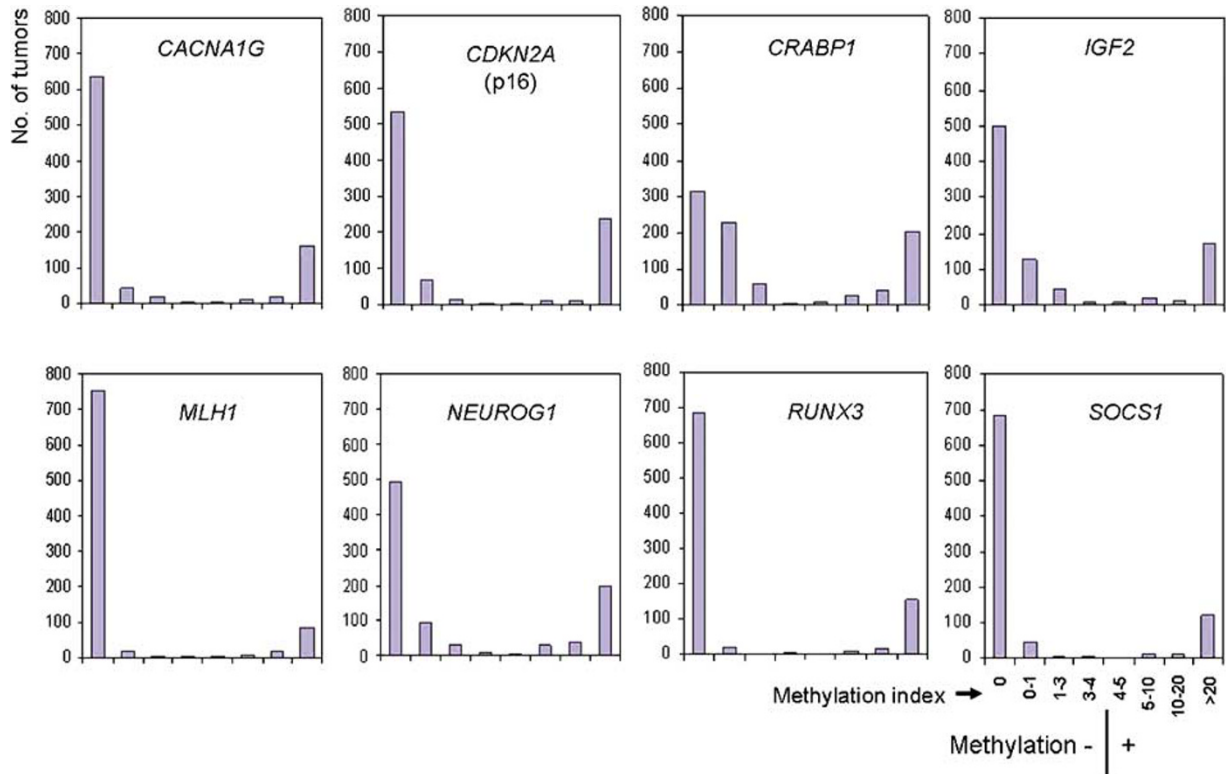
### Overall Distribution of Methylation Index Values in Each Locus

Utilizing MethyLight technology, we quantified DNA methylation in a panel of eight promoters (*CACNA1G*, *CDKN2A*, *CRABP1*, *IGF2*, *MLH1*, *NEUROG1*, *RUNX3* and *SOCS1*).<sup>6,10</sup> All of the eight markers have been shown to be sensitive and specific for CIMP-high, and the use of the eight markers as a CIMP diagnostic panel has been validated.<sup>10</sup> Figure 1 shows a distribution of methylation index values at each locus in the 889 colorectal cancers. Only rare cases showed methylation index values between 3 and 5 in any locus, indicating that methylation status (either negative, methylation index  $< 4$  or positive, methylation index  $> 4$ ) at each locus was unequivocally determined in a vast majority of cases. There was no tumor with methylation index = 4, 10 or 20, exactly, at any locus.

### Classification of CIMP-High, CIMP-Low and CIMP-0 Tumors

The CpG island methylator phenotype-high (CIMP-high) was defined as the presence of  $\geq 6/8$  methylated (positive) loci, CIMP-low as the presence of 1/8 to 5/8 methylated loci, and CIMP-0 as the absence (0/8) of methylated loci based on the data that CIMP-high and CIMP-low are associated with *BRAF* mutation and *KRAS* mutation, respectively.<sup>10,17</sup> Among the 889 colorectal cancers, there were 133 (15%) CIMP-high tumors, 340 (38%) CIMP-low tumors and 416 (47%) CIMP-0 tumors.

We selected tumors with at least 40% tumor cellularity for this study, and also assessed the



**Figure 1** Distribution of methylation index (percentage of methylated reference) values in 889 colorectal cancers.

average tumor cellularity in each CIMP category, to exclude the possibility that differences in methylation index values might be simply caused by differences in tumor cellularity. The average tumor cellularities were 54.8% in CIMP-high, 56.2% in CIMP-low and 57.7% in CIMP-0 tumors, and there were no significant differences (CIMP-high vs CIMP-low,  $P=0.45$ ; CIMP-low vs CIMP-0,  $P=0.27$ ).

**Low-Level Methylation at an Individual Promoter was Common in CIMP-Low Tumors**

In addition to methylation index cutoff of 4 that has been repeatedly validated as a cutoff for biologically significant methylation,<sup>10,24,32,33</sup> we assigned additional arbitrary cutoffs of low-level methylation, methylation index <10 and <20. We tested these two additional cutoffs (instead of one additional cutoff), because these cutoffs were arbitrary. First, we examined the frequency of methylation index >0 and <20 (also written as methylation index 0–20) in each promoter among CIMP-high tumors, CIMP-low tumors and CIMP-0 tumors (Table 1). The overall frequency of methylation index 0–20 in each promoter was almost always higher in CIMP-low tumors than in CIMP-high and CIMP-0 tumors (with the only exception of a higher frequency of *MLH1* methylation index 0–20 in CIMP-high tumors than CIMP-low tumors).

Considering biological significance of the methylation index cutoff of 4 (ie methylation positive vs negative), we next analyzed the frequencies of methylation index >0 and <4 (between 0 and 4) (also written as methylation index 0–4) and methylation index 4–20, separately. To compare the frequencies of methylation index 0–4 in *CACNA1G*, for example, we limited the denominator to tumors with *CACNA1G* methylation index <4 (methylation negative) (Table 1). This was because, compared to CIMP-0 tumors, CIMP-high and CIMP-low tumors had a higher proportion of tumors positive for *CACNA1G* methylation (methylation index >4), which, if included in the denominator, would have by itself decreased the frequency of methylation index 0–4. Compared to CIMP-0 tumors, CIMP-low tumors consistently showed higher frequencies of methylation index 0–4 in all eight loci ( $P=0.004$ ), and the differences were highly significant in *CRABP1*, *MLH1*, *NEUROG1* and *RUNX3* (all  $P<0.001$ ).

To compare the frequencies of methylation index 4–10 (or 4–20) in *CACNA1G*, for example, we limited the denominator to tumors with *CACNA1G* methylation index >4 (methylation positive) (Table 1). This was because, compared to CIMP-high tumors, CIMP-low tumors had a higher proportion of tumors negative for *CACNA1G* methylation (methylation index <4), which, if included in the denominator, would have by itself decreased

**Table 1** Frequency of low methylation index at each CIMP-high-specific promoter among CIMP-high, CIMP-low and CIMP-0 colorectal cancers

| Methylation index at each locus     | CIMP-high (N = 133) | P-value (CIMP-high vs low) | CIMP-low (N = 340) | P-value (CIMP-0 vs low) | CIMP-0 (N = 416) |
|-------------------------------------|---------------------|----------------------------|--------------------|-------------------------|------------------|
| <b>CACNA1G</b>                      |                     |                            |                    |                         |                  |
| >0 and <20                          | 11 (8.3%)           | 0.01                       | 58 (17%)           | <0.0001                 | 27 (6.5%)        |
| All <4 (all negative cases)         | —                   |                            | 278                |                         | 416              |
| >0 and <4 (% in all negative cases) | —                   |                            | 33 (12%)           | 0.01                    | 27 (6.5%)        |
| All >4 (all positive cases)         | 130                 |                            | 62                 |                         | —                |
| 4–10 (% in all positive cases)      | 3 (2.3%)            | 0.0001                     | 12 (19%)           |                         | —                |
| 4–20 (% in all positive cases)      | 8 (6.2%)            | <0.0001                    | 25 (40%)           |                         | —                |
| <b>CDKN2A</b>                       |                     |                            |                    |                         |                  |
| >0 and <20                          | 10 (7.5%)           | 0.01                       | 55 (16%)           | 0.18                    | 53 (13%)         |
| All <4 (all negative cases)         | —                   |                            | 195                |                         | 416              |
| >0 and <4 (% in all negative cases) | —                   |                            | 32 (16%)           | 0.22                    | 53 (13%)         |
| All >4 (all positive cases)         | 121                 |                            | 145                |                         | —                |
| 4–10 (% in all positive cases)      | 6 (5.0%)            | 0.46                       | 11 (7.6%)          |                         | —                |
| 4–20 (% in all positive cases)      | 6 (5.0%)            | 0.005                      | 23 (16%)           |                         | —                |
| <b>CRABP1</b>                       |                     |                            |                    |                         |                  |
| >0 and <20                          | 13 (9.8%)           | <0.0001                    | 176 (52%)          | 0.02                    | 179 (43%)        |
| All <4 (all negative cases)         | —                   |                            | 190                |                         | 416              |
| >0 and <4 (% in all negative cases) | —                   |                            | 111 (58%)          | 0.0004                  | 179 (43%)        |
| All >4 (all positive cases)         | 132                 |                            | 150                |                         | —                |
| 4–10 (% in all positive cases)      | 1 (0.8%)            | <0.0001                    | 35 (23%)           |                         | —                |
| 4–20 (% in all positive cases)      | 12 (9.1%)           | <0.0001                    | 65 (43%)           |                         | —                |
| <b>IGF2</b>                         |                     |                            |                    |                         |                  |
| >0 and <20                          | 5 (3.8%)            | <0.0001                    | 109 (32%)          | 0.04                    | 105 (25%)        |
| All <4 (all negative cases)         | —                   |                            | 259                |                         | 419              |
| >0 and <4 (% in all negative cases) | —                   |                            | 74 (29%)           | 0.34                    | 105 (25%)        |
| All >4 (all positive cases)         | 130                 |                            | 81                 |                         | —                |
| 4–10 (% in all positive cases)      | 2 (1.5%)            | <0.0001                    | 24 (30%)           |                         | —                |
| 4–20 (% in all positive cases)      | 3 (2.3%)            | <0.0001                    | 35 (43%)           |                         | —                |
| <b>MLH1</b>                         |                     |                            |                    |                         |                  |
| >0 and <20                          | 24 (18%)            | <0.0001                    | 24 (7.1%)          | <0.0001                 | 3 (0.7%)         |
| All <4 (all negative cases)         | —                   |                            | 323                |                         | 416              |
| >0 and <4 (% in all negative cases) | —                   |                            | 18 (5.6%)          | <0.0001                 | 3 (0.7%)         |
| All >4 (all positive cases)         | 96                  |                            | 17                 |                         | —                |
| 4–10 (% in all positive cases)      | 8 (8.3%)            | 1.0                        | 1 (5.9%)           |                         | —                |
| 4–20 (% in all positive cases)      | 21 (22%)            | 0.23                       | 6 (35%)            |                         | —                |
| <b>NEUROG1</b>                      |                     |                            |                    |                         |                  |
| >0 and <20                          | 18 (14%)            | <0.0001                    | 119 (35%)          | <0.0001                 | 63 (15%)         |
| All <4 (all negative cases)         | —                   |                            | 205                |                         | 416              |
| >0 and <4 (% in all negative cases) | —                   |                            | 66 (32%)           | <0.0001                 | 63 (15%)         |
| All >4 (all positive cases)         | 130                 |                            | 135                |                         | —                |
| 4–10 (% in all positive cases)      | 9 (6.9%)            | 0.01                       | 23 (17%)           |                         | —                |
| 4–20 (% in all positive cases)      | 16 (12%)            | <0.0001                    | 53 (39%)           |                         | —                |
| <b>RUNX3</b>                        |                     |                            |                    |                         |                  |
| >0 and <20                          | 8 (6.0%)            | 0.12                       | 37 (11%)           | <0.0001                 | 7 (1.7%)         |
| All <4 (all negative cases)         | —                   |                            | 290                |                         | 416              |
| >0 and <4 (% in all negative cases) | —                   |                            | 19 (6.6%)          | 0.0009                  | 7 (1.7%)         |
| All >4 (all positive cases)         | 130                 |                            | 50                 |                         | —                |
| 4–10 (% in all positive cases)      | 3 (2.3%)            | 0.002                      | 8 (16%)            |                         | —                |
| 4–20 (% in all positive cases)      | 8 (6.2%)            | <0.0001                    | 18 (36%)           |                         | —                |
| <b>SOCS1</b>                        |                     |                            |                    |                         |                  |
| >0 and <20                          | 20 (15%)            | 0.29                       | 39 (11%)           | 0.01                    | 26 (6.3%)        |
| All <4 (all negative cases)         | —                   |                            | 285                |                         | 416              |
| >0 and <4 (% in all negative cases) | —                   |                            | 23 (8.1%)          | 0.35                    | 26 (6.3%)        |
| All >4 (all positive cases)         | 95                  |                            | 54                 |                         | —                |
| 4–10 (% in all positive cases)      | 6 (6.3%)            | 0.03                       | 10 (18%)           |                         | —                |
| 4–20 (% in all positive cases)      | 14 (15%)            | 0.03                       | 16 (29%)           |                         | —                |

CIMP, CpG island methylator phenotype.

the frequency of methylation index 4–10 (or 4–20). Compared to CIMP-high tumors, CIMP-low tumors consistently showed higher frequencies of methylation index 4–20 in all eight loci ( $P=0.004$ ) and methylation index 4–10 in all eight but one locus (*MLH1*) ( $P=0.03$ ) (Table 1). The differences were statistically significant in many individual loci (*CACNA1G*, *CRABP1*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1*, all  $P\leq 0.03$ ) regardless of the cutoff used (methylation index 4–10 or 4–20).

### DNA Methylation Levels at CpG Islands not Included in the CIMP Panel

In addition to the eight promoter CpG islands in the CIMP panel, we examined DNA methylation levels in six other CpG islands (*CHFR*, *HIC1*, *IGFBP3*, *MGMT*, *MINT31* and *WRN*), which were not in the CIMP panel (Table 2). Among 888 tumors, sensitivity and specificity of each marker for the CIMP-high diagnosis were as follows: *CHFR*, 90 and 65%; *HIC1*, 89 and 55%; *IGFBP3*, 57 and 78%; *MGMT*, 50 and 65%; *MINT31*, 97 and 78%; and *WRN*, 94 and 73%, respectively. Sensitivities or specificities of these six markers for CIMP-high were considerably lower than those of the eight CIMP panel markers.<sup>10</sup>

With regard to *WRN* methylation, low methylation index was significantly more common in CIMP-low tumors than in CIMP-high and CIMP-0 tumors (Table 2). However, for the other five markers, the relationship between low methylation index and CIMP-low was not persistent. Our results suggest that the relationship between CIMP-low tumors and low methylation index is mainly limited to CpG islands that are highly correlated with CIMP-high.

### Prevalences (Frequencies) of Methylation Index 0–4, 4–10 and 10–20 Across 8 CIMP-Specific Promoters in Each Tumor

We examined an overall prevalence of low-level methylation in the eight CIMP panel markers (in addition to low-level methylation in each marker as in Table 1). Table 3 illustrates methods to assess overall prevalence (frequency) of methylation index 0–4 (highlighted by blue) in each tumor, using examples of tumors. The frequency of loci with methylation index 0–4 in each tumor was calculated as ‘(the number of loci with methylation index 0–4)/(the number of all negative loci with methylation index <4).’ We limited the denominator to negative loci with methylation index <4; otherwise, the presence of more positive loci by itself would have made methylation index 0–4 less prevalent. When there were more positive loci (methylation index >4) in a given tumor (eg case ID 6–9 in Table 3), there were fewer negative loci with methylation index <4.

The frequency of loci with methylation index 4–10 (highlighted by yellow in Table 3) in each

tumor was calculated as (the number of loci with methylation index 4–10)/(the number of all positive loci with methylation index >4). Likewise, the frequency of loci with methylation index 10–20 (highlighted by red in Table 3) in each tumor was calculated as ‘(the number of loci with methylation index 10–20)/(the number of all positive loci with methylation index >4).’ We limited the denominator to positive loci with methylation index >4; otherwise, the presence of fewer positive loci by itself would have made methylation index 4–10 (or methylation index 10–20) less prevalent.

### Low Methylation Index 4–10 or 4–20 is More Common in CIMP-Low Tumors than in CIMP-High Tumors

Table 4 shows a distribution of the frequencies of methylation index 4–10, according to the number of positive loci. For tumors with a given number of positive loci, an average of the frequencies of methylation index 4–10 was calculated as ‘ $\sum(\text{frequency of methylation index 4–10})/(\text{the number of cases})$ .’ The average reflected how often a tumor with a given number of positive loci showed methylation index 4–10 among all positive loci (methylation index >4). The average for CIMP-low tumors (19.4%) was much higher than that for CIMP-high tumors (4.2%,  $P=0.002$ ) (Table 4), indicating that low-level methylation (methylation index 4–10) was far more common in CIMP-low tumors than CIMP-high tumors.

Similarly, Table 5 shows a distribution of the frequencies of methylation index 4–20, according to the number of positive loci. For tumors with a given number of positive loci, an average of the frequencies of methylation index 4–20 was calculated as ‘ $\sum(\text{frequency of methylation index 4–20})/(\text{the number of cases})$ .’ The average for CIMP-low tumors (36.8%) was much higher than that for CIMP-high tumors (9.4%,  $P<0.0001$ ) (Table 5), indicating that methylation index 4–20 was far more common in CIMP-low tumors than in CIMP-high tumors.

We also calculated a weighted average of the frequencies of low-level methylation as ‘ $\sum\{(\text{frequency of methylation index 4–10}) + (\text{frequency of methylation index 10–20})/2\}/(\text{number of cases})$ ,’ to put twice more weights on methylation index 4–10 than methylation index 10–20 (Table 5). The weighted average for CIMP-low tumors (28.1%) was much higher than that for CIMP-high tumors (6.8%,  $P<0.0001$ ), further supporting that low-level methylation was more common in CIMP-low tumors than in CIMP-high tumors.

### Low Methylation Index 0–4 is More Common in CIMP-Low Tumors than in CIMP-0 Tumors

Table 6 shows a distribution of the frequencies of methylation index 0–4, according to the number of negative loci. For tumors with a given number of negative loci, an average of the frequencies

**Table 2** Frequency of low methylation index at each non-CIMP-specific locus among CIMP-high, CIMP-low and CIMP-0 tumors

| Methylation index at each locus     | CIMP-high<br>(N = 133) | P-value<br>(CIMP-high<br>vs low) | CIMP-low<br>(N = 339) | P-value<br>(CIMP-0<br>vs low) | CIMP-0<br>(N = 416) |
|-------------------------------------|------------------------|----------------------------------|-----------------------|-------------------------------|---------------------|
| <i>CHFR</i>                         |                        |                                  |                       |                               |                     |
| >0 and <20                          | 32 (24%)               | 0.32                             | 97 (29%)              | 0.02                          | 88 (21%)            |
| All <4 (all negative cases)         | —                      |                                  | 177                   |                               | 315                 |
| >0 and <4 (% in all negative cases) | —                      |                                  | 33 (19%)              | 0.03                          | 36 (11%)            |
| All >4 (all positive cases)         | 120                    |                                  | 162                   |                               | —                   |
| 4–10 (% in all positive cases)      | 5 (4.2%)               | 0.02                             | 20 (12%)              |                               | —                   |
| 4–20 (% in all positive cases)      | 25 (21%)               | 0.0009                           | 64 (40%)              |                               | —                   |
| <i>HIC1</i>                         |                        |                                  |                       |                               |                     |
| >0 and <20                          | 16 (12%)               | <0.0001                          | 140 (41%)             | <0.0001                       | 233 (56%)           |
| All <4 (all negative cases)         | —                      |                                  | 130                   |                               | 243                 |
| >0 and <4 (% in all negative cases) | —                      |                                  | 78 (60%)              | 0.57                          | 153 (63%)           |
| All >4 (all positive cases)         | 118                    |                                  | 209                   |                               | —                   |
| 4–10 (% in all positive cases)      | 2 (1.7%)               | <0.0001                          | 37 (18%)              |                               | —                   |
| 4–20 (% in all positive cases)      | 6 (5.1%)               | <0.0001                          | 62 (30%)              |                               | —                   |
| <i>IGFBP3</i>                       |                        |                                  |                       |                               |                     |
| >0 and <20                          | 52 (39.1%)             | 0.57                             | 123 (36%)             | 0.0006                        | 103 (25%)           |
| All <4 (all negative cases)         | —                      |                                  | 220                   |                               | 357                 |
| >0 and <4 (% in all negative cases) | —                      |                                  | 58 (26%)              | 0.08                          | 72 (20%)            |
| All >4 (all positive cases)         | 81                     |                                  | 119                   |                               | —                   |
| 4–10 (% in all positive cases)      | 11 (14%)               | 0.08                             | 28 (24%)              |                               | —                   |
| 4–20 (% in all positive cases)      | 41 (51%)               | 0.58                             | 65 (55%)              |                               | —                   |
| <i>MGMT</i>                         |                        |                                  |                       |                               |                     |
| >0 and <20                          | 18 (14%)               | 0.10                             | 29 (8.5%)             | 0.10                          | 51 (12%)            |
| All <4 (all negative cases)         | —                      |                                  | 192                   |                               | 288                 |
| >0 and <4 (% in all negative cases) | —                      |                                  | 15 (7.8%)             | 0.33                          | 30 (10%)            |
| All >4 (all positive cases)         | 166                    |                                  | 148                   |                               | —                   |
| 4–10 (% in all positive cases)      | 3 (1.8%)               | 0.48                             | 5 (3.4%)              |                               | —                   |
| 4–20 (% in all positive cases)      | 9 (5.4%)               | 0.20                             | 14 (9.5%)             |                               | —                   |
| <i>MINT31</i>                       |                        |                                  |                       |                               |                     |
| >0 and <20                          | 3 (2.3%)               | <0.0001                          | 159 (47%)             | 0.56                          | 204 (49%)           |
| All <4 (all negative cases)         | —                      |                                  | 209                   |                               | 369                 |
| >0 and <4 (% in all negative cases) | —                      |                                  | 116 (56%)             | 0.21                          | 185 (52%)           |
| All >4 (all positive cases)         | 131                    |                                  | 131                   |                               | —                   |
| 4–10 (% in all positive cases)      | 1 (0.8%)               | <0.0001                          | 24 (18%)              |                               | —                   |
| 4–20 (% in all positive cases)      | 2 (1.5%)               | <0.0001                          | 43 (33%)              |                               | —                   |
| <i>WRN</i>                          |                        |                                  |                       |                               |                     |
| >0 and <20                          | 10 (7.5%)              | <0.0001                          | 118 (35%)             | <0.0001                       | 89 (21%)            |
| All <4 (all negative cases)         | —                      |                                  | 169                   |                               | 364                 |
| >0 and <4 (% in all negative cases) | —                      |                                  | 64 (38%)              | <0.0001                       | 66 (18%)            |
| All >4 (all positive cases)         | 127                    |                                  | 170                   |                               | —                   |
| 4–10 (% in all positive cases)      | 4 (3.1%)               | 0.001                            | 24 (14%)              |                               | —                   |
| 4–20 (% in all positive cases)      | 8 (6.3%)               | <0.0001                          | 54 (32%)              |                               | —                   |

CIMP, CpG island methylator phenotype.

of methylation index 0–4 was calculated as ‘ $\sum$ (frequency of methylation index 0–4)/(the number of cases)’. The average reflected how often a tumor with a given number of negative loci showed methylation index 0–4 among all negative loci (methylation index <4). The average for CIMP-low tumors (18.8%) was significantly higher than that for CIMP-0 tumors (13.9%,  $P < 0.0001$ ) (Table 6), indicating that low-level CpG island methylation (methylation index 0–4) was more common in CIMP-low tumors than CIMP-0 tumors.

The frequencies of methylation index 0–4 in CIMP-high tumors were quite unevenly distributed,

because of fewer (only 1–2) negative loci (methylation index <4) in CIMP-high tumors (Table 6). Therefore, it was difficult to draw conclusion from a comparison of the frequencies of methylation index 0–4 between CIMP-high and CIMP-low.

## Discussion

The purpose of this study is to examine density of CpG island methylation in CIMP-low and CIMP-high colorectal cancers, using a large population-based sample ( $N = 889$ ). We analyzed 14 CpG

**Table 3** Example of methylation index values and methods to calculate the frequency of low methylation index (>0 and <20) in an individual tumor

| Case ID | Methylation index (percentage of methylated reference) |                     |               |             |             |                |              |              | CIMP status* | No. of positive loci (M index >4) | No. of negative loci (M index <4) | Frequency of loci with M index 0-4 among all negative loci (M index <4) | Frequency of loci with M index 4-10 among all positive loci (M index >4) | Frequency of loci with M index 10-20 among all positive loci (M index >4) |
|---------|--------------------------------------------------------|---------------------|---------------|-------------|-------------|----------------|--------------|--------------|--------------|-----------------------------------|-----------------------------------|-------------------------------------------------------------------------|--------------------------------------------------------------------------|---------------------------------------------------------------------------|
|         | <i>CACNA1G</i>                                         | <i>CDKN2A (p16)</i> | <i>CRABP1</i> | <i>IGF2</i> | <i>MLH1</i> | <i>NEUROG1</i> | <i>RUNX3</i> | <i>SOCS1</i> |              |                                   |                                   |                                                                         |                                                                          |                                                                           |
| 1       | 0                                                      | 0                   | 1.6           | 0           | 0           | 0              | 0            | 0            | 0            | 8                                 | 1/8 (13%)                         | -                                                                       | -                                                                        |                                                                           |
| 2       | 0                                                      | 7.7                 | 0             | 0           | 0           | 0.8            | 0            | 0            | Low          | 1                                 | 7                                 | 1/7 (14%)                                                               | 1/1 (100%)                                                               | 0/1 (0%)                                                                  |
| 3       | 0                                                      | 2.1                 | 69.3          | 12.4        | 0           | 1.4            | 0            | 0            | Low          | 2                                 | 6                                 | 2/6 (33%)                                                               | 0/2 (0%)                                                                 | 1/2 (50%)                                                                 |
| 4       | 8.2                                                    | 0.7                 | 89.0          | 0           | 0           | 0              | 16.9         | 0            | Low          | 3                                 | 5                                 | 1/5 (20%)                                                               | 1/3 (33%)                                                                | 1/3 (33%)                                                                 |
| 5       | 16.3                                                   | 16.0                | 3.1           | 58.7        | 3.1         | 2.4            | 11.9         | 0            | Low          | 4                                 | 4                                 | 3/4 (75%)                                                               | 0/4 (0%)                                                                 | 3/4 (75%)                                                                 |
| 6       | 7.7                                                    | 93.9                | 12.3          | 9.0         | 0.5         | 11.1           | 0            | 0            | Low          | 5                                 | 3                                 | 1/3 (33%)                                                               | 2/5 (40%)                                                                | 2/5 (40%)                                                                 |
| 7       | 47.0                                                   | 62.2                | 15.5          | 76.6        | 0           | 3.5            | 5.6          | 5.9          | High         | 6                                 | 2                                 | 1/2 (50%)                                                               | 2/6 (33%)                                                                | 1/6 (17%)                                                                 |
| 8       | 19.8                                                   | 123.4               | 68.6          | 56.8        | 0           | 6.4            | 89.1         | 13.0         | High         | 7                                 | 1                                 | 0/1 (0%)                                                                | 1/7 (14%)                                                                | 2/7 (29%)                                                                 |
| 9       | 102.2                                                  | 7.0                 | 16.0          | 68.6        | 10.5        | 5.4            | 28.4         | 17.9         | High         | 8                                 | 0                                 | -                                                                       | 2/8 (25%)                                                                | 3/8 (38%)                                                                 |

CIMP, CpG island methylator phenotype; M index, methylation index (percentage of methylated reference). For this table, cases with low-level methylation were particularly selected for the purpose of illustrating the methods to calculate the frequency of low-level methylation. Methylation index 0–4 is highlighted by blue, methylation index 4–10 by yellow and methylation index 10–20 by red. \*CIMP-0 is defined as the absence of positive (methylation index >4) loci, CIMP-low as 1–5 positive loci, and CIMP-high as 6–8 positive loci.

islands including 8 promoters (*CACNA1G*, *CDKN2A (p16)*, *CRABP1*, *IGF2*, *MLH1*, *NEUROG1*, *RUNX3* and *SOCS1*), all of which have been shown to be specifically methylated in CIMP-high tumors.<sup>6,10</sup> The other six CpG islands (*CHFR*, *HIC1*, *IGFBP3*, *MGMT*, *MINT31* and *WRN*) were not as sensitive or specific for CIMP-high. We showed that low-level methylation at the eight CIMP-high-specific loci were more common in CIMP-low tumors than in CIMP-high and CIMP-0 tumors, while there were no consistent differences in methylation levels at the other six non-CIMP-specific loci between CIMP-low, CIMP-high and CIMP-0. Our findings may provide evidence for molecular differences between CIMP-low, CIMP-high and CIMP-0 tumors. CIMP-low tumors tend to methylate CIMP-specific CpG islands at low levels, and some of methylated CpG islands may cross the threshold of gene silencing. We have previously shown that low-level methylation (methylation index <4) is not associated with gene silencing, whereas methylation index >4 is associated with gene silencing.<sup>18,24</sup> In contrast, CIMP-high tumors tend to methylate most CIMP-specific CpG islands at high levels. Our data may provide evidence for a difference in pathogenesis of DNA methylation between CIMP-low and CIMP-high tumors.

Transcriptional inactivation of tumor suppressor genes by promoter CpG island methylation is an important mechanism in human carcinogenesis.<sup>2,34–38</sup> To measure the degree of DNA methylation, we used quantitative PCR assays (MethyLight),<sup>20</sup> which is essential to reproducibly differentiate low-level methylation from high-level methylation.<sup>21,24</sup> Our resource of a large number of samples of colorectal cancer (relatively unbiased

samples compared to retrospective or single-hospital-based samples), derived from two large prospective cohorts, has enabled us to precisely estimate the frequency of low-level CpG island methylation in colorectal cancers at a population level. Our data also provide a rationale to use quantitative DNA methylation analysis in research and clinical settings. If we had interpreted low-level methylation (methylation index 0–4) as methylation positive, which might occur in sensitive nonquantitative methylation-specific PCR assays, we might have misdiagnose a considerable number of CIMP-low tumors as CIMP-high tumors. Thus, quantitative assessment of DNA methylation is very important.<sup>24,25,27,39,40</sup>

Our findings provide additional supporting evidence for molecular features of CIMP-low tumors different from CIMP-high and CIMP-0 tumors. CIMP-low tumors generally appear to methylate at multiple CIMP-specific CpG islands at low levels, whereas CIMP-0 tumors tend to show more loci with no evidence of methylation and CIMP-high tumors tend to show more loci with high-level methylation. We have previously shown that CIMP-low is associated with male sex and *KRAS* mutations, while CIMP-high is associated with female sex and *BRAF* mutations and CIMP-0 is associated with wild-type *KRAS/BRAF* genes and shows no sex predilection.<sup>17</sup> We have recently shown a possible link between CIMP-low, microsatellite instability-low, *MGMT* methylation and *KRAS* mutation, supporting unique molecular features of CIMP-low.<sup>18</sup> Together with our current data, CIMP-low appears to be a different phenotype from CIMP-high and CIMP-0. Further studies are necessary to identify the best set of markers for the diagnosis of



**Table 4** Overall frequency of low methylation index 4–10 among all positive loci (methylation index >4) according to the number of methylation-positive loci

| No. of positive loci<br>(M index > 4) | Total N | Frequency of methylation index 4–10 among all positive loci (with methylation index > 4) |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | Average frequency (%) |                 |   |    |       |
|---------------------------------------|---------|------------------------------------------------------------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-----------------------|-----------------|---|----|-------|
|                                       |         | 0%<br>(0 loci)                                                                           | 1/8 loci<br>(13%) | 1/7 loci<br>(14%) | 1/6 loci<br>(17%) | 1/5 loci<br>(20%) | 1/4 loci<br>(25%) | 2/7 loci<br>(29%) | 1/3 loci<br>(33%) | 2/5 loci<br>(40%) | 1/2 loci<br>(50%) | 3/5 loci<br>(60%) | 2/3 loci<br>(67%) | 3/4 loci<br>(75%) |                       | 100% (all loci) |   |    |       |
| 1                                     | 163     | 129                                                                                      |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                       |                 |   | 34 | 20.9  |
| 2                                     | 85      | 56                                                                                       |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | 25                |                       |                 |   | 4  | 19.4  |
| 3                                     | 32      | 14                                                                                       |                   |                   |                   |                   |                   |                   |                   |                   | 15                |                   |                   |                   |                       |                 | 2 | 1  | 22.9  |
| 4                                     | 34      | 19                                                                                       |                   |                   |                   |                   |                   | 12                |                   |                   |                   |                   |                   | 1                 |                       |                 |   | 2  | 14.7  |
| 5                                     | 26      | 16                                                                                       |                   |                   |                   |                   | 6                 |                   |                   |                   |                   |                   | 3                 |                   |                       | 1               |   |    | 11.5  |
| 6                                     | 27      | 16                                                                                       |                   |                   | 7                 |                   |                   |                   |                   |                   | 4                 |                   |                   |                   |                       |                 |   |    | 9.3   |
| 7                                     | 46      | 37                                                                                       |                   | 8                 |                   |                   |                   |                   |                   | 1                 |                   |                   |                   |                   |                       |                 |   |    | 3.1   |
| 8                                     | 60      | 50                                                                                       | 7                 |                   |                   |                   |                   |                   | 3                 |                   |                   |                   |                   |                   |                       |                 |   |    | 2.7   |
| CIMP-low (1–5)                        | 340     | 234                                                                                      |                   |                   |                   |                   | 6                 | 12                |                   |                   | 15                | 3                 | 26                | 1                 | 2                     | 2               |   | 39 | 19.4* |
| CIMP-high (6–8)                       | 133     | 103                                                                                      | 7                 | 8                 | 7                 |                   |                   | 3                 |                   | 1                 | 4                 |                   |                   |                   |                       |                 |   |    | 4.2*  |

CIMP, CpG island methylator phenotype; M index, methylation index (percentage of methylated reference).

CIMP-0 tumors are not listed because there was no positive locus.

\*P = 0.002 (CIMP-low vs CIMP-high) by the Wilcoxon's rank sum test.

**Table 5** Overall frequency of low methylation index 4–20 among all positive loci (with methylation index >4) according to the number of methylation-positive loci

| No. of positive loci<br>(M index > 4) | Total N | Frequency of methylation index 4–20 among all positive loci (with methylation index > 4) |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | Average frequency (%) | Weighted average* (%) |                    |
|---------------------------------------|---------|------------------------------------------------------------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-----------------------|-----------------------|--------------------|
|                                       |         | 0%<br>(0 loci)                                                                           | 1/8 loci<br>(13%) | 1/7 loci<br>(14%) | 1/6 loci<br>(17%) | 1/5 loci<br>(20%) | 1/4 loci<br>(25%) | 2/7 loci<br>(29%) | 1/3 loci<br>(33%) | 3/8 loci<br>(38%) | 2/5 loci<br>(40%) | 3/7 loci<br>(43%) | 1/2 loci<br>(50%) | 3/5 loci<br>(60%) | 5/8 loci<br>(63%) | 2/3 loci<br>(67%) | 3/4 loci<br>(75%) | 4/5 loci<br>(80%) |                       |                       | 100%<br>(all loci) |
| 1                                     | 163     | 103                                                                                      |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   |                   | 60                    | 36.8                  | 28.8               |
| 2                                     | 85      | 29                                                                                       |                   |                   |                   |                   |                   |                   |                   |                   |                   | 40                |                   |                   |                   |                   |                   |                   | 16                    | 42.4                  | 30.9               |
| 3                                     | 32      | 7                                                                                        |                   |                   |                   |                   |                   | 13                |                   |                   |                   |                   |                   |                   | 8                 |                   |                   |                   | 4                     | 42.7                  | 32.8               |
| 4                                     | 34      | 12                                                                                       |                   |                   |                   | 9                 |                   |                   |                   |                   |                   | 10                |                   |                   |                   | 1                 |                   |                   | 2                     | 29.4                  | 22.1               |
| 5                                     | 26      | 10                                                                                       |                   |                   | 7                 |                   |                   |                   |                   | 7                 |                   |                   |                   | 1                 |                   |                   |                   | 1                 |                       | 21.5                  | 16.5               |
| 6                                     | 27      | 11                                                                                       |                   |                   | 10                |                   |                   | 5                 |                   |                   |                   | 1                 |                   |                   |                   |                   |                   |                   |                       | 14.2                  | 11.7               |
| 7                                     | 46      | 26                                                                                       |                   | 15                |                   |                   |                   |                   |                   |                   | 3                 |                   |                   |                   |                   |                   |                   |                   |                       | 8.7                   | 5.9                |
| 8                                     | 60      | 37                                                                                       | 16                |                   |                   | 2                 |                   |                   | 4                 |                   |                   |                   |                   |                   | 1                 |                   |                   |                   |                       | 7.7                   | 5.2                |
| CIMP-low (1–5)                        | 340     | 161                                                                                      |                   |                   |                   | 7                 | 9                 | 13                |                   | 7                 |                   | 50                | 1                 |                   | 8                 | 1                 | 1                 | 82                | 36.8^                 | 28.1#                 |                    |
| CIMP-high (6–8)                       | 133     | 74                                                                                       | 16                | 15                | 10                | 2                 | 2                 | 5                 | 4                 | 3                 | 1                 | 1                 |                   | 1                 |                   |                   |                   |                   |                       | 9.4^                  | 6.8#               |

CIMP, CpG island methylator phenotype; M index, methylation index (percentage of methylated reference).

CIMP-0 tumors are not listed because there was no positive locus.

<sup>a</sup>Weighted average was calculated as  $\frac{[\sum((\text{frequency of methylation index 4–10})+(\text{frequency of methylation index 10–20})/2)]}{(\text{number of cases})}$ , to put twice more weights on methylation index 4–10 than methylation index 10–20; ^P < 0.0001 (CIMP-low vs CIMP-high) by the Wilcoxon's rank sum test; #P < 0.0001 (CIMP-low vs CIMP-high) by the Wilcoxon's rank sum test.

**Table 6** Overall frequency of low methylation index 0–4 among all negative loci (methylation index <4) according to the number of methylation-positive loci

| No. of positive loci (M index >4) | No. of negative loci (M index <4) | Total N | Frequency of methylation index > 0 and < 4 among all negative loci (methylation index <4) |                |                |                |                |                |                |                | Average frequency (%) |                |                |                |                |                |                |                |                |                |                |                |                |
|-----------------------------------|-----------------------------------|---------|-------------------------------------------------------------------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|                                   |                                   |         | 0% (0 loci)                                                                               | 1/8 loci (13%) | 1/7 loci (14%) | 1/6 loci (17%) | 1/5 loci (20%) | 1/4 loci (25%) | 2/7 loci (29%) | 1/3 loci (33%) |                       | 3/8 loci (38%) | 2/5 loci (40%) | 3/7 loci (43%) | 1/2 loci (50%) | 4/7 loci (57%) | 3/5 loci (60%) | 5/8 loci (63%) | 2/3 loci (67%) | 5/7 loci (71%) | 3/4 loci (75%) | 4/5 loci (80%) | 5/6 loci (83%) |
| 0 (CIMP-0)                        | 8                                 | 416     | 175                                                                                       | 108            | 50             | 27             | 11             | 14             | 30             | 19             | 36                    | 20             | 6              | 3              | 23             | 3              | 1              | 1              | 1              | 1              | 1              | 1              | 13.9*          |
| 1                                 | 7                                 | 163     | 53                                                                                        | 33             | 11             | 11             | 14             | 10             | 7              | 7              | 7                     | 3              | 3              | 6              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 18.2           |
| 2                                 | 6                                 | 85      | 33                                                                                        | 11             | 11             | 11             | 14             | 10             | 7              | 7              | 7                     | 3              | 3              | 6              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 17.1           |
| 3                                 | 5                                 | 32      | 11                                                                                        | 11             | 11             | 11             | 14             | 10             | 7              | 7              | 7                     | 3              | 3              | 6              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 21.9           |
| 4                                 | 4                                 | 34      | 12                                                                                        | 14             | 14             | 14             | 14             | 10             | 7              | 7              | 7                     | 3              | 3              | 6              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 23.5           |
| 5                                 | 3                                 | 26      | 14                                                                                        | 14             | 14             | 14             | 14             | 10             | 7              | 7              | 7                     | 3              | 3              | 6              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 2              | 18.0           |
| 6                                 | 2                                 | 27      | 15                                                                                        | 15             | 15             | 15             | 15             | 15             | 15             | 15             | 15                    | 15             | 15             | 15             | 15             | 15             | 15             | 15             | 15             | 15             | 15             | 15             | 22.2           |
| 7                                 | 1                                 | 46      | 37                                                                                        | 37             | 37             | 37             | 37             | 37             | 37             | 37             | 37                    | 37             | 37             | 37             | 37             | 37             | 37             | 37             | 37             | 37             | 37             | 37             | 19.6           |
| CIMP-low (1–5)                    | 3–7                               | 340     | 123                                                                                       | 123            | 123            | 123            | 123            | 123            | 123            | 123            | 123                   | 123            | 123            | 123            | 123            | 123            | 123            | 123            | 123            | 123            | 123            | 123            | 18.8*          |
| CIMP-high (6–7)                   | 1–2                               | 73      | 52                                                                                        | 52             | 52             | 52             | 52             | 52             | 52             | 52             | 52                    | 52             | 52             | 52             | 52             | 52             | 52             | 52             | 52             | 52             | 52             | 52             | 20.5           |

CIMP, CpG island methylator phenotype; M index, methylation index. CIMP-high tumors with all 8 methylated promoters are not listed because there was no negative locus. \*P < 0.0001 (CIMP-0 vs CIMP-low) by the Wilcoxon's rank sum test.

CIMP-low, because the eight markers we used are sensitive and specific for CIMP-high,<sup>6,10</sup> but not intended to separate CIMP-low from CIMP-0.

Pathogenesis of CIMP-high and CIMP-low tumors are still elusive.<sup>16</sup> Unlike microsatellite instability in which we have known a role of the mismatch repair system, the etiologies of CIMP as well as high and low-level methylation may not be as straightforward. Considering our current data, we can speculate that CIMP-high tumors may be caused by a fundamental defect in the methyl-group metabolism, resulting in high-level methylation. In contrast, CIMP-low tumors may have either a more subtle defect, or may be a deviation of the CIMP-0 pathway. Further studies are necessary to elucidate pathogenesis of various forms of CIMP in colorectal cancer.

In conclusion, CIMP-low tumors show not only few methylated CIMP-specific CpG islands, but also frequent low-level DNA methylation in individual CpG islands, in contrast to CIMP-high and CIMP-0 tumors. Our data may raise a possibility for a difference in pathogenesis of DNA methylation between CIMP-low and CIMP-high tumors. Further studies are necessary to elucidate the mechanisms of low-level and high-level methylation in CIMP-low and CIMP-high tumors, respectively.

### Acknowledgements

This work was supported by National Institute of Health (NIH) Grants P01 CA87969 and P01 CA55075, and in part by the Bennett Family Fund for Targeted Therapies Research, and by the Entertainment Industry Foundation (EIF) through the EIF National Colorectal Cancer Research Alliance (NCCRA). We deeply thank the Nurses' Health Study and Health Professionals Follow-up Study cohort participants who have generously agreed to provide us with biological specimens and information through responses to questionnaires. We thank Graham Colditz, Walter Willett and many other staff members who implemented and have maintained the cohort studies. No conflict of interest is present.

### References

- Issa JP. CpG island methylator phenotype in cancer. *Nat Rev Cancer* 2004;4:988–993.
- Agrawal A, Murphy RF, Agrawal DK. DNA methylation in breast and colorectal cancers. *Mod Pathol* 2007;20:711–721.
- Toyota M, Ahuja N, Ohe-Toyota M, *et al*. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci USA* 1999;96:8681–8686.
- Hawkins N, Norrie M, Cheong K, *et al*. CpG island methylation in sporadic colorectal cancers and its relationship to microsatellite instability. *Gastroenterology* 2002;122:1376–1387.

- 5 Kambara T, Simms LA, Whitehall VLJ, *et al*. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. *Gut* 2004;53:1137–1144.
- 6 Weisenberger DJ, Siegmund KD, Campan M, *et al*. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet* 2006;38:787–793.
- 7 Samowitz W, Albertsen H, Herrick J, *et al*. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology* 2005;129:837–845.
- 8 Ogino S, Cantor M, Kawasaki T, *et al*. CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. *Gut* 2006;55:1000–1006.
- 9 Kawasaki T, Noshio K, Ohnishi M, *et al*. Correlation of beta-catenin localization with cyclooxygenase-2 expression and CpG island methylator phenotype (CIMP) in colorectal cancer. *Neoplasia* 2007;9:569–577.
- 10 Ogino S, Kawasaki T, Kirkner GJ, *et al*. Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. *J Mol Diagn* 2007;9:305–314.
- 11 Van Rijnsoever M, Elsaleh H, Joseph D, *et al*. CpG island methylator phenotype is an independent predictor of survival benefit from 5-fluorouracil in stage III colorectal cancer. *Clin Cancer Res* 2003;9:2898–2903.
- 12 Ward RL, Cheong K, Ku SL, *et al*. Adverse prognostic effect of methylation in colorectal cancer is reversed by microsatellite instability. *J Clin Oncol* 2003;21:3729–3736.
- 13 Ogino S, Meyerhardt JA, Kawasaki T, *et al*. CpG island methylation, response to combination chemotherapy, and patient survival in advanced microsatellite stable colorectal carcinoma. *Virchows Arch* 2007;450:529–537.
- 14 Shen L, Catalano PJ, Benson III AB, *et al*. Association between DNA methylation and shortened survival in patients with advanced colorectal cancer treated with 5-fluorouracil based chemotherapy. *Clin Cancer Res* 2007;13:6093–6098.
- 15 Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* 2007;50:113–130.
- 16 Grady WM. CIMP and colon cancer gets more complicated. *Gut* 2007;56:1498–1500.
- 17 Ogino S, Kawasaki T, Kirkner GJ, *et al*. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. *J Mol Diagn* 2006;8:582–588.
- 18 Ogino S, Kawasaki T, Kirkner GJ, *et al*. Molecular correlates with MGMT promoter methylation and silencing support CpG island methylator phenotype-low (CIMP-low) in colorectal cancer. *Gut* 2007;56:1409–1416.
- 19 Ogino S, Kawasaki T, Kirkner GJ, *et al*. 18q loss of heterozygosity in microsatellite stable colorectal cancer is correlated with CpG island methylator phenotype-negative (CIMP-0) and inversely with CIMP-low and CIMP-high. *BMC Cancer* 2007;7:72.
- 20 Eads CA, Danenberg KD, Kawakami K, *et al*. MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res* 2000;28:E32.
- 21 Lehmann U, Kreipe H. Real-time PCR-based assay for quantitative determination of methylation status. *Methods Mol Biol* 2004;287:207–218.
- 22 Chen J, Rocken C, Lofton-Day C, *et al*. Molecular analysis of APC promoter methylation and protein expression in colorectal cancer metastasis. *Carcinogenesis* 2005;26:37–43.
- 23 Harden SV, Guo Z, Epstein JI, *et al*. Quantitative GSTP1 methylation clearly distinguishes benign prostatic tissue and limited prostate adenocarcinoma. *J Urol* 2003;169:1138–1142.
- 24 Ogino S, Kawasaki T, Brahmandam M, *et al*. Precision and performance characteristics of bisulfite conversion and real-time PCR (MethyLight) for quantitative DNA methylation analysis. *J Mol Diagn* 2006;8:209–217.
- 25 Lehmann U, Wingen LU, Brakensiek K, *et al*. Epigenetic defects of hepatocellular carcinoma are already found in non-neoplastic liver cells from patients with hereditary haemochromatosis. *Hum Mol Genet* 2007;16:1335–1342.
- 26 Kawakami K, Ruzkiewicz A, Bennett G, *et al*. DNA hypermethylation in the normal colonic mucosa of patients with colorectal cancer. *Br J Cancer* 2006;94:593–598.
- 27 Ebert MP, Mooney SH, Tonnes-Priddy L, *et al*. Hypermethylation of the TPEF/HPP1 gene in primary and metastatic colorectal cancers. *Neoplasia* 2005;7:771–778.
- 28 Colditz GA, Hankinson SE. The nurses' health study: lifestyle and health among women. *Nat Rev Cancer* 2005;5:388–396.
- 29 Giovannucci E, Liu Y, Rimm EB, *et al*. Prospective study of predictors of vitamin D status and cancer incidence and mortality in men. *J Natl Cancer Inst* 2006;98:451–459.
- 30 Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. *N Engl J Med* 2007;356:2131–2142.
- 31 Ogino S, Kawasaki T, Brahmandam M, *et al*. Sensitive sequencing method for KRAS mutation detection by Pyrosequencing. *J Mol Diagn* 2005;7:413–421.
- 32 Eads CA, Lord RV, Kurumboor SK, *et al*. Fields of aberrant CpG island hypermethylation in Barrett's esophagus and associated adenocarcinoma. *Cancer Res* 2000;60:5021–5026.
- 33 Widschwendter M, Siegmund KD, Muller HM, *et al*. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. *Cancer Res* 2004;64:3807–3813.
- 34 Laird PW. Cancer epigenetics. *Hum Mol Genet* 2005;14 Spec No 1:R65–R76.
- 35 Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006;7:21–33.
- 36 Dong SM, Lee EJ, Jeon ES, *et al*. Progressive methylation during the serrated neoplasia pathway of the colorectum. *Mod Pathol* 2005;18:170–178.
- 37 Kawaguchi K, Oda Y, Saito T, *et al*. DNA hypermethylation status of multiple genes in soft tissue sarcomas. *Mod Pathol* 2006;19:106–114.
- 38 Choi IS, Estecio MR, Nagano Y, *et al*. Hypomethylation of LINE-1 and Alu in well-differentiated neuroendocrine tumors (pancreatic endocrine tumors and carcinoid tumors). *Mod Pathol* 2007;20:802–810.
- 39 Shaw RJ, Akufu-Tetteh EK, Risk JM, *et al*. Methylation enrichment pyrosequencing: combining the specificity of MSP with validation by pyrosequencing. *Nucleic Acids Res* 2006;34:e78.
- 40 Ehrich M, Nelson MR, Stanssens P, *et al*. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci USA* 2005;102:15785–15790.