protein-coding genes are targets of DNA methylation in BCa, and their urinary methylation appears to be a useful biomarker [28,29]. For instance, methylation of 11 protein-coding genes found in urine sediments revealed the presence of BCa with a high sensitivity and specificity [30], and in another study a panel of three genes (growth differentiation factor 15 [*GDF15*], transmembrane protein with EGF-like and two follistatin-like domains 2 [*TMEFF2*], and vimentin [*VIM*]) in urine could be used to accurately detect BCa [31]. In the present study, we show for the first time that methylation of miRNA genes could serve as a biomarker for detection of BCa. Methylation of miRNA genes was readily detectable in voided urine from cancer patients, and its levels were dramatically reduced after tumor resection, confirming its tumor specificity. We also showed that a combination of multiple miRNA genes could accurately distinguish between preoperative and postoperative urine samples.

Our study has several limitations. The prognostic value of miRNA gene methylation remains unclear, because the prognosis of the patients in this study is not yet available. A follow-up study in post-treatment patients will be needed to test whether urinary methylation can predict outcome or detect BCa recurrence. In addition, urinary methylation in non-BCa patients (eg, patients with other types of cancer) should be tested to evaluate the specificity of our method. Further studies to address these issues would contribute to overcoming the difficulties in translating our present findings into clinical practice.

5. Conclusions

We identified four miRNA genes that are frequent targets of epigenetic silencing in BCa. Although their specific functions in bladder carcinogenesis remain unknown, it is evident that restoration of these miRNAs may be an effective anticancer therapy. Furthermore, methylation of these miRNA genes in urine specimens could serve as a useful and noninvasive biomarker for accurate detection of BCa.

Author contributions: Hiromu Suzuki had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Suzuki, Toyota.
Acquisition of data: Shimizu, Ashida, Hatahira, Yamamoto, Maruyama, Kai.
Analysis and interpretation of data: Shimizu, Suzuki, Nojima.
Drafting of the manuscript: Shimizu, Suzuki.
Critical revision of the manuscript for important intellectual content: Taiji Tsukamoto.
Statistical analysis: Nojima.
Obtaining funding: Suzuki, Toyota, Tsukamoto.
Administrative, technical, or material support: Kitamura, Masumori, Tokino, Imai, Tsukamoto.
Supervision: Minoru Toyota, Taiji Tsukamoto.
Other (specify): None.

Financial disclosures: Hiromu Suzuki certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Funding/Support and role of the sponsor: This study was supported in part by a Program for Developing a Supporting System for Upgrading Education and Research from the Ministry of Education, Culture, Sports, Science, and Technology (T. Tsukamoto, M. Toyota); a Grant-in-Aid for the Third-term Comprehensive 10-year Strategy of Cancer Control from the Ministry of Health, Labor, and Welfare, Japan (M. Toyota, H. Suzuki); the A3 Foresight Program from the Japan Society for Promotion of Science (H. Suzuki); and the Project for Developing Innovative Research on Cancer Therapeutics (P-DIRECT; H. Suzuki).

Acknowledgement statement: The authors thank the staff of the Departments of Urology at Sapporo Medical University Hospital and NTT East Sapporo Hospital for their kind assistance in the collection of specimens and Dr William Goldman for editing the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2012.11.030.

References

[1] He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 2004;5:522–31.
[2] Croce CM. Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet 2009;10:704–14.
[3] Chen Q, Chen X, Zhang M, Fan Q, Luo S, Cao X. miR-137 is frequently down-regulated in gastric cancer and is a negative regulator of Cdc42. Dig Dis Sci 2011;56:2009–16.
[4] Catto JW, Miah S, Owen HC, et al. Distinct microRNA alterations characterize high- and low-grade bladder cancer. Cancer Res 2009;69:8472–81.
[5] Neely LA, Rieger-Christ KM, Neto BS, et al. A microRNA expression ratio defining the invasive phenotype in bladder tumors. Urol Oncol 2010;28:39–48.
[6] Adam L, Zhong M, Choi W, et al. miR-200 expression regulates epithelial-to-mesenchymal transition in bladder cancer cells and reverses resistance to epidermal growth factor receptor therapy. Clin Cancer Res 2009;15:5060–72.
[7] Saito Y, Liang G, Egger G, et al. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell 2006;9:435–43.
[8] Suzuki H, Takatsuka S, Akashi H, et al. Genome-wide profiling of chromatin signatures reveals epigenetic regulation of microRNA genes in colorectal cancer. Cancer Res 2011;71:5646–58.
[9] Suzuki H, Yamamoto E, Nojima M, et al. Methylation-associated silencing of microRNA-34b/c in gastric cancer and its involvement in an epigenetic field defect. Carcinogenesis 2010;31:2066–73.
[10] Maruyama R, Suzuki H, Yamamoto E, Imai K, Shinomura Y. Emerging links between epigenetic alterations and dysregulation of non-coding RNAs in cancer. Tumour Biol 2012;33:277–85.
[11] Esteller M. Non-coding RNAs in human disease. Nat Rev Genet 2011;12:861–74.
[12] Wiklund ED, Bramsen JB, Hulf T, et al. Coordinated epigenetic repression of the miR-200 family and miR-205 in invasive bladder cancer. Int J Cancer 2011;128:1327–34.
[13] Dudziec E, Miah S, Choudhry HM, et al. Hypermethylation of CpG islands and shores around specific microRNAs and mirtrons is associated with the phenotype and presence of bladder cancer. Clin Cancer Res 2011;17:1287–96.

[14] Balaguer F, Link A, Lozano JJ, et al. Epigenetic silencing of miR-137 is an early event in colorectal carcinogenesis. Cancer Res 2010;70:6609–18.
[15] Liu M, Lang N, Qiu M, et al. miR-137 targets Cdc42 expression, induces cell cycle G1 arrest and inhibits invasion in colorectal cancer cells. Int J Cancer 2011;128:1269–79.
[16] Kozaki K, Imoto I, Mogi S, Omura K, Inazawa J. Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. Cancer Res 2008;68:2094–105.
[17] Szulwach KE, Li X, Smrt RD, et al. Cross talk between microRNA and epigenetic regulation in adult neurogenesis. J Cell Biol 2010;189:127–41.
[18] Smrt RD, Szulwach KE, Pfeiffer RL, et al. MicroRNA miR-137 regulates neuronal maturation by targeting ubiquitin ligase mind bomb-1. Stem Cells 2010;28:1060–70.
[19] Lujambio A, Ropero S, Ballestar E, et al. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. Cancer Res 2007;67:1424–9.
[20] Ando T, Yoshida T, Enomoto S, et al. DNA methylation of microRNA genes in gastric mucosae of gastric cancer patients: its possible involvement in the formation of epigenetic field defect. Int J Cancer 2009;124:2367–74.
[21] Agirre X, Vilas-Zornoza A, Jimenez-Velasco A, et al. Epigenetic silencing of the tumor suppressor microRNA Hsa-miR-124a regulates CDK6 expression and confers a poor prognosis in acute lymphoblastic leukemia. Cancer Res 2009;69:4443–53.
[22] Furuta M, Kozaki KI, Tanaka S, Arii S, Imoto I, Inazawa J. miR-124 and miR-203 are epigenetically silenced tumor-suppressive microRNAs in hepatocellular carcinoma. Carcinogenesis 2010;31:766–76.
[23] Lujambio A, Calin GA, Villanueva A, et al. A microRNA DNA methylation signature for human cancer metastasis. Proc Natl Acad Sci U S A 2008;105:13556–61.
[24] Lodygin D, Tarasov V, Epanchintsev A, et al. Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer. Cell Cycle 2008;7:2591–600.
[25] Hildebrandt MA, Gu J, Lin J, et al. Hsa-miR-9 methylation status is associated with cancer development and metastatic recurrence in patients with clear cell renal cell carcinoma. Oncogene 2010;29:5724–8.
[26] Yamada Y, Enokida H, Kojima S, et al. MiR-96 and miR-183 detection in urine serve as potential tumor markers of urothelial carcinoma: correlation with stage and grade, and comparison with urinary cytology. Cancer Sci 2011;102:522–9.
[27] Miah S, Dudziec E, Drayton RM, et al. An evaluation of urinary microRNA reveals a high sensitivity for bladder cancer. Br J Cancer 2012;107:123–8.
[28] Nishiyama N, Arai E, Chihara Y, et al. Genome-wide DNA methylation profiles in urothelial carcinomas and urothelia at the precancerous stage. Cancer Sci 2010;101:231–40.
[29] Cairns P. Gene methylation and early detection of genitourinary cancer: the road ahead. Nat Rev Cancer 2007;7:531–43.
[30] Yu J, Zhu T, Wang Z, et al. A novel set of DNA methylation markers in urine sediments for sensitive/specific detection of bladder cancer. Clin Cancer Res 2007;13:7296–304.
[31] Costa VL, Henrique R, Danielsen SA, et al. Three epigenetic biomarkers, GDF15, TMEFF2, and VIM, accurately predict bladder cancer from DNA-based analyses of urine samples. Clin Cancer Res 2010;16:5842–51.

GENES, CHROMOSOMES & CANCER 52:140–149 (2013)

Association Between Genomic Alterations and Metastatic Behavior of Colorectal Cancer Identified by Array-Based Comparative Genomic Hybridization

Takeshi Sawada,[1,2] Eiichiro Yamamoto,[3] Hiromu Suzuki,[1]* Masanori Nojima,[4] Reo Maruyama,[1] Yoshihiro Shioi,[5] Risaburo Akasaka,[5] Seiko Kamimae,[1] Taku Harada,[1] Masami Ashida,[1] Masahiro Kai,[1] Yasushi Adachi,[3] Hiroyuki Yamamoto,[3] Kohzoh Imai,[6] Minoru Toyota,[1] Fumio Itoh,[2] and Tamotsu Sugai[5]*

[1]Department of Molecular Biology, Sapporo Medical University, Sapporo, Japan
[2]Division of Gastroenterology and Hepatology, St Marianna University School of Medicine, Kawasaki, Japan
[3]First Department of Internal Medicine, Sapporo Medical University, Sapporo, Japan
[4]Department of Public Health, Sapporo Medical University, Sapporo, Japan
[5]Division of Molecular Diagnostic Pathology, Department of Pathology, School of Medicine, Iwate Medical University, Morioka, Japan
[6]Division of Novel Therapy for Cancer, The Advanced Clinical Research Center, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Colorectal cancers (CRCs) exhibit multiple genetic alterations, including allelic imbalances (copy number alterations, CNAs) at various chromosomal loci. In addition to genetic aberrations, DNA methylation also plays important roles in the development of CRC. To better understand the clinical relevance of these genetic and epigenetic abnormalities in CRC, we performed an integrative analysis of copy number changes on a genome-wide scale and assessed mutations of *TP53*, *KRAS*, *BRAF*, and *PIK3CA* and DNA methylation of six marker genes in single glands isolated from 39 primary tumors. Array-based comparative genomic hybridization (array-CGH) analysis revealed that genomic losses commonly occurred at 3q26.1, 4q13.2, 6q21.32, 7q34, 8p12-23.3, 15qcen and 18, while gains were commonly found at 1q21.3-23.1, 7p22.3-q34, 13q12.11-14.11, and 20. The total numbers and lengths of the CNAs were significantly associated with the aberrant DNA methylation and Dukes' stages. Moreover, hierarchical clustering analysis of the array-CGH data suggested that tumors could be categorized into four subgroups. Tumors with frequent DNA methylation were most strongly enriched in subgroups with infrequent CNAs. Importantly, Dukes' D tumors were enriched in the subgroup showing the greatest genomic losses, whereas Dukes' C tumors were enriched in the subgroup with the greatest genomic gains. Our data suggest an inverse relationship between chromosomal instability and aberrant methylation and a positive association between genomic losses and distant metastasis and between genomic gains and lymph node metastasis in CRC. Therefore, DNA copy number profiles may be predictive of the metastatic behavior of CRCs.

INTRODUCTION

Colorectal cancers (CRCs) develop through multiple genetic alterations, including allelic losses at chromosomal loci (e.g., 5q, 17p, and 18q) (Fearon and Vogelstein, 1990). In addition, epigenetic changes, including aberrant DNA methylation and histone modifications, are also strongly implicated in the pathogenesis of CRC, and a subset of CRCs show concurrent hypermethylation in multiple loci, which is now classified as the CpG island methylator phenotype (CIMP) (Toyota et al., 1999). Recent studies have shown that there are two types of CRCs with distinct genomic abnormalities: chromosomal instability (CIN), which accounts for 80–85% of sporadic CRCs and was originally characterized based on the presence of aneuploid/polyploid karyotypes, and microsatellite instability (MSI), also termed MIN, which accounts for 15–20% of sporadic CRCs and is characterized by mismatch repair defects and a near-diploid karyotype (Grady and Carethers, 2008). CIN cancers exhibit gains and losses at multiple chromosomal loci (copy number alterations; CNAs) (Rajagopalan and Lengauer, 2004), whereas MSI cancers show considerable

Additional Supporting Information may be found in the online version of this article.

Supported by Japan Society for Promotion of Science (Grant-in-Aid for Scientific Research (C) [H.S.]); a Grant-in-Aid for the Third-term Comprehensive 10-year Strategy for Cancer Control (M.T. and H.S.); Ministry of Health, Labor, and Welfare, Japan (Grant-in-Aid for Cancer Research [M.T. and H.S.)], Japan Society for Promotion of Science (A3 foresight program [H.S.)].

*Correspondence to: Tamotsu Sugai, Division of Molecular Diagnostic Pathology, Department of Pathology, School of Medicine, Iwate Medical University, 19-1, Morioka 020-8505, Japan. E-mail: tsugai@cocoa.ocn.ne.jp or Hiromu Suzuki, Department of Molecular Biology, Sapporo Medical University, S1, W17, Chuo-ku, Sapporo 060-8556, Japan. E-mail: hsuzuki@sapmed.ac.jp

Received 21 February 2012; Accepted 22 August 2012

DOI 10.1002/gcc.22013

Published online 17 October 2012 in Wiley Online Library (wileyonlinelibrary.com).

overlap with CIMP cancers (Toyota et al., 1999; Ogino et al., 2006; Weisenberger et al., 2006).

In addition to the commonly observed CRC-related allelic losses on chromosome arms 5q, 17p, and 18q, gains and losses on many other chromosomes have been identified using conventional comparative genomic hybridization (CGH) analysis (Ried et al., 1996; Meijer et al., 1998). Diep et al. (2006) conducted a meta-analysis of the chromosomal changes in a series of 859 CRC specimens identified using CGH and reported that specific CNAs are associated with each step during the progression of CRC. Still, conventional CGH has limited resolution and can only detect CNAs of ~10 Mb or greater in length. On the other hand, array-based CGH (array-CGH) can detect genetic changes with a resolution of 1 Mb or less, making it a powerful tool with which to analyze genomic alterations (Douglas et al., 2004; Jones et al., 2005).

From a clinical viewpoint, previous studies have shown that CRCs can be categorized into distinct subgroups based on the characteristics of their CNAs (Hermsen et al., 2002), and such subtyping has predictive value with respect to prognosis (Poulogiannis et al., 2010) and the response to chemotherapy (Postma et al., 2009). Similarly, a number of studies have shown that epigenetic alterations, especially CIMP, are strongly associated with the clinical behavior of CRCs (Shen et al., 2007a; Jover et al., 2011). However, although it is recognized that CRCs develop via multiple molecular pathways, including CIN, MSI, and CIMP (Jass, 2007; Shen et al., 2007b; Issa, 2008; Hinoue et al., 2012), the associations between genetic and epigenetic abnormalities are still not fully understood. In this study, we performed an integrative analysis of copy number changes on genome-wide scale and assessed genetic mutation of *TP53*, *KRAS*, *BRAF*, and *PIK3CA* and DNA methylation of six marker genes within crypts isolated from surgically resected CRCs, and assessed their relevance to the clinicopathological characteristics.

MATERIALS AND METHODS

Patients and Tissue Samples

A total of 39 primary CRCs and corresponding normal tissue specimens were obtained from consecutive patients at the Iwate Medical University Hospital. Informed consent was obtained from all patients before collection of the specimens, and approval of this study was obtained from the Institutional Review Board of Iwate Medical University. The clinicopathological features of the patients are summarized in Table 1. Pathological diagnosis and staging were performed using a combination of the Japanese classification (Japanese Society for Cancer of the Colon and Rectum, 1997) and modified Dukes' classification (Turnbull et al., 1967). Tumor locations were classified as left- or right-sided and rectal.

TABLE 1. Clinicopathological Features of the CRC Samples Used in this Study

Age (years, median ± SD)	69 ± 11.7
Sex	
Male	24 (62%)
Female	15 (38%)
Location	
Right	14 (36%)
Left	5 (13%)
Rectum	20 (51%)
Histology	
Mod	29 (74%)
Well	7 (17%)
Pap	1 (3%)
Por	1 (3%)
Muc	1 (3%)
Dukes' stage	
A	7 (18%)
B	5 (13%)
C	13 (33%)
D	14 (36%)
Lymph node metastasis	
Positive	23 (59%)
Negative	16 (41%)

Isolation of Glands and Genomic DNA Extraction

Glands were isolated from the tumors and normal mucosae as described previously (Arai and Kino, 1989; Nakamura et al., 1994). The isolated glands were routinely processed to confirm their nature using paraffin-embedded histological sections. Contamination by other materials such as interstitial cells was not evident in the samples examined, which is consistent with previous reports (Sugai et al., 2000; Sugai et al., 2005). Genomic DNA was extracted using the standard phenol-chloroform procedure.

Analysis of *TP53*, *BRAF*, *KRAS*, and *PIK3CA* Mutations

Exons 5–8 of *TP53* were PCR amplified and then analyzed using single-strand conformational polymorphism (SSCP). PCR amplification, PCR-SSCP, and the sequencing of *TP53* were performed as described previously (Dix et al., 1994; Habano et al., 1996; Sugai et al., 2000). In addition,

mutation of codon 600 of *BRAF* and codons 12 and 13 of *KRAS* was examined by pyrosequencing using *BRAF* and *KRAS* pyro kits (Qiagen) according to the manufacturer's instructions, and exons 9 and 20 of *PIK3CA* were directly sequenced as described previously (Jhawer et al., 2008).

DNA Methylation Analysis

CpG island methylation was analyzed as described previously (Toyota et al., 2008). Briefly, genomic DNA (1 μg) was modified with sodium bisulfite using an EpiTect Bisulfite Kit (Qiagen). Pyrosequencing was carried out using a PSQ 96MA system (Qiagen) with a Pyro Gold Reagent Kit (Qiagen), and the results were analyzed using Pyro Q-CpG software (Qiagen). A cutoff value of 15% was used to define genes as methylation-positive. Tumors were defined as "tumors with frequent DNA methylation" when methylation was detected in three or more loci out of six markers (*MINT1*, *MINT2*, *MINT12*, *MINT31*, *CDKN2A*, and *MLH1*).

Array-Based CGH

Array CGH analysis was performed as described previously (Igarashi et al., 2010). Briefly, 500 ng of genomic DNA and gender-matched reference DNA (Promega) were digested with *Alu*I and *Rsa*I before labeling and hybridization. Using a Genomic DNA Enzymatic Labeling Kit (Agilent Technologies), tumor DNA and reference DNA were, respectively, labeled with Cy5 and Cy3 after being hybridized to a Human Genome CGH Microarray Kit 105A (G4412A; Agilent Technologies), which contains approximately 99,000 probes annotated against National Center for Biotechnology Information Build 36. The ADM-2 algorithm included in the Genomic Workbench software ver. 5 (Agilent Technologies) was used to identify DNA copy number aberrations. A copy number loss was defined as a log2 ratio <-0.5, and a copy number gain was defined as a log2 ratio >0.5. All genomic positions were defined according to the University of California Santa Cruz Human version hg 18. Unsupervised hierarchical analysis was performed on the log2 Cy5/Cy3 fluorescence ratio data using the Ward's linkage algorithm (JMP version 8, SAS Institute, Cary, NC).

Statistical Analysis

Continuous data was analyzed using *t* tests (for two groups) or ANOVA with a post hoc Tukey's HSD test (for more than two groups). To detect specific differences within groups, adjusted standardized residuals were calculated for the categorical data. If the absolute values of the residuals were more than 1.96, we considered them significantly different from a random distribution. *P* values <0.05 were considered significant. All statistical analyses were performed using SPSS 20 (IBM Corporation, Somers, NY) and Prism 5 (GraphPad Software, La Jolla, CA).

RESULTS

Overview of Array-CGH Analysis

The results of our array-CGH analysis of crypts obtained from 39 CRC tumors are summarized in Figure 1. Genomic losses were commonly observed at several loci, including 3q26.1 (75%), 4q13.2 (80%), 6q21.32 (83%), 7q34 (58%), 8p12-23.3 (55%), 15qcen (50%), and 18 (80%), while gains were commonly observed at 1q21.3-23.1 (41%), 7p22.3-q34 (48%), 13q12.11-14.11 (50%), and 20 (75%). Large genomic losses (>10 Mb in length) were frequently seen at 8p (54%), 18p (59%), and 18q (77%), and large gains (>10 Mb in length) were seen at 7 (41%), 13q (46%), 20p (44%), and 20q (74%) (Fig. 1, Supporting Information Fig. 1). These findings are mostly consistent with earlier results obtained using conventional CGH and array-CGH (Ried et al., 1996; Meijer et al., 1998; Douglas et al., 2004; Jones et al., 2005; Diep et al., 2006).

Mutation and Methylation Analysis

Among the 39 CRC specimens tested, *TP53* and *KRAS* mutations were found in 21 (54%) and 15 (38%), respectively, which is also consistent with earlier findings (Supporting Information Fig. 1 and Table 1) (Dix et al., 1994; Smith et al., 2002; Baldus et al., 2010). However, the frequency of samples with mutations in both *TP53* and *KRAS* (15%) was higher than previously reported (9%) (Smith et al., 2002), probably because our study included tumors at more advanced stages. By contrast, *PIK3CA* mutation was found in only four (10%) specimens, which is less frequent than previously reported (Samuels et al., 2004; Baldus et al., 2010). *BRAF* mutation was not detected in any samples.

Bisulfite-pyrosequencing analysis revealed that 9 of the 39 tumors (23%) exhibited methylation at 3 or more loci, although none showed methylation of *MLH1*. *KRAS* mutation was more prevalent among tumors with frequent DNA methylation (6/9, 67%) than among those without frequent methylation (9/30, 30%). Previous studies demonstrated that

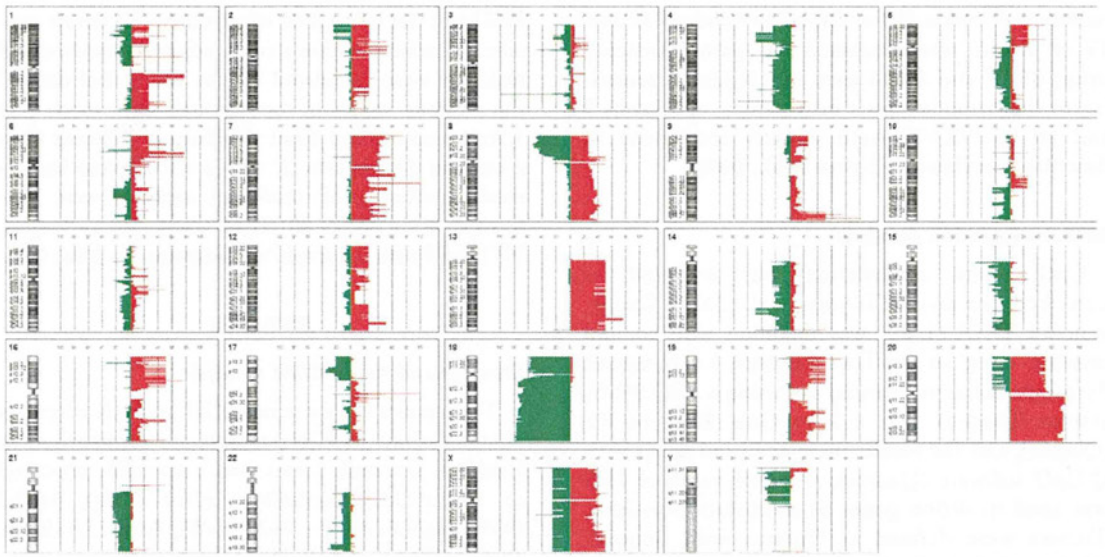

Figure 1. Summary of chromosomal aberrations and their frequencies in 39 CRC specimens determined using array-CGH analysis. Losses (green bars) are displayed on the left, and gains (red bars) are on the right. The chromosome ideogram was generated using Genomic Workbench software.

BRAF mutation and MSI are significantly more prevalent among CIMP-high (CIMP-H or CIMP1) CRCs, whereas *KRAS* mutation is more prevalent among CIMP-low (CIMP-L or CIMP2) tumors (Ogino et al., 2006; Shen et al., 2007b; Hinoue et al., 2012).

Inverse Correlation Between DNA Methylation and Chromosomal Alterations

To quantitatively evaluate copy number aberrations on a genome-wide scale, we calculated the total numbers and lengths of CNAs (losses + gains) identified by the array-CGH analysis. We observed a strong correlation between the total numbers of CNAs in the CRC samples tested and the total lengths of the CNAs (Supporting Information Fig. 2). We therefore used the total CNA length as an index representing the degree of chromosomal alteration and assessed the relationship between total CNA length and methylation status. We found that total CNA lengths were smaller in CRCs with frequent DNA methylation than in those without frequent methylation ($P = 0.033$, Fig. 2C). Interestingly, when we analyzed genomic losses and gains separately, we again observed a significant difference in the magnitude of the losses between frequent methylation-positive and -negative tumors, whereas no such difference was found for gains (Figs. 2A and 2B). We also analyzed the relationship between CNAs and mutation of *TP53* or *KRAS* but found no statistically significant correlations (Supporting Information Fig. 3).

Chromosomal Alterations and Their Association with Clinical Stage in CRC

To determine whether chromosomal alterations accumulate during the progression of CRCs, we assessed the CNA status of tumors at each Dukes' stage. When genomic gains and losses were analyzed separately, we found a tendency toward greater genomic losses in tumors at higher Dukes' stages, but the trend was not statistically significant (Fig. 3A). By contrast, Dukes' C tumors showed the greatest genomic gains, whereas Dukes' D tumors exhibited unexpectedly small gains (Fig. 3B). The total CNA lengths were greatest in Dukes' C tumors, and again Dukes' D tumors showed less chromosomal alteration than Dukes' C tumors (Fig. 3C).

We found that losses at several loci, including 3p24.3, 4p13-15.31, 5qcen-11.2, 8p11-q11, 9p21.3-21.1, 17q24.2-24.3, and 22q13.31, were prevalent in tumors with distant metastasis (Dukes' D) (Supporting Information Table 2), while gains at 11q13.1-13.2, 17q12, and 17q21.2 were prevalent among Dukes' D tumors (Supporting Information Table 3). In addition, losses at 4q21-34, 5q12.1, 5q35.3, and 11qcen-12.1 and gains at 2p15-16.1, 2p13.3, 5p13, 5p35.3, 6p21, 8q12.1-12.3, and 19q13.31 were commonly observed in tumors with lymph node metastasis (Supporting Information Tables 4, 5).

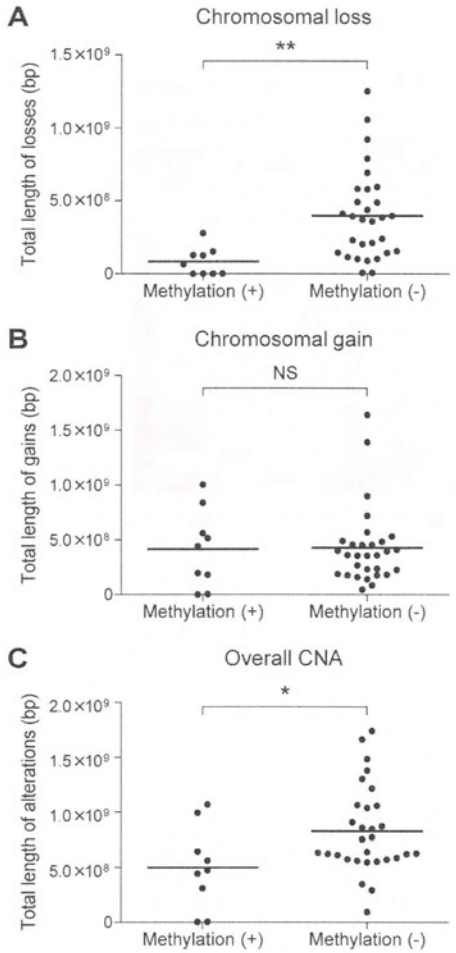

Figure 2. Association between CNAs and DNA methylation status. Total lengths of genomic losses (A) and gains (B) and overall CNAs (C) in tumors with and without frequent DNA methylation are shown. Each dot represents a single tumor. Methylation (+), tumors with frequent DNA methylation; Methylation (−), tumors without frequent DNA methylation; $*P < 0.05$; $**P < 0.01$; NS, not significant.

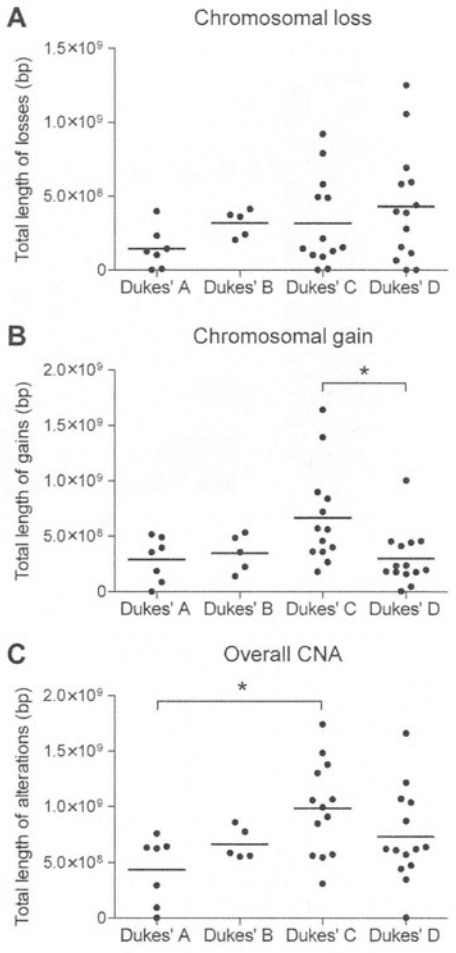

Figure 3. Association between CNAs and Dukes' stages. Total lengths of genomic losses (A), gains (B), and overall CNAs (C) in CRCs at each Dukes' stage are shown. $*P < 0.05$.

Clustering Analysis of CNAs and Their Association with Clinical Stage in CRC

Previous studies have shown that categorization of CRCs according to their chromosomal aberrations has strong relevance to their clinical behavior (Hermsen et al., 2002; Postma et al., 2009; Poulogiannis et al., 2010). For that reason, we carried out unsupervised clustering analysis using our array-CGH data (excluding the sex chromosomes) and then compared the results with genetic mutations and epigenetic alterations (Fig. 4A). We found that CRCs could be subcategorized into at least four clusters based on their CNAs. Gene mutations, DNA methylation status, and genomic alterations on representative chromosomes in each cluster are summarized in Table 2.

Tumors in cluster 1 are characterized by infrequent genomic losses and gains (Figs. 4A–4C). Losses were most prevalent among tumors in cluster 2, while gains were most prevalent among tumors in cluster 4 (Figs. 4A–4C). The total CNA lengths were greater in tumors in clusters 2 and 4 than in clusters 1 and 3 (Fig. 4D). Tumors with frequent DNA methylation were most strongly enriched in cluster 1 (5 of 9, 56%). Tumors in cluster 1 were also characterized by frequent *KRAS* mutation (6 of 10, 60%) and infrequent *p53* mutation (3 of 10, 30%), whereas *p53* mutation was most prevalent in cluster 3 tumors (10 of 15, 67%), although the difference was not statistically significant. Importantly, Dukes' D tumors were highly enriched in cluster 2, within which tumors showed the greatest genomic losses. By contrast, Dukes' C tumors were enriched in cluster 4 and

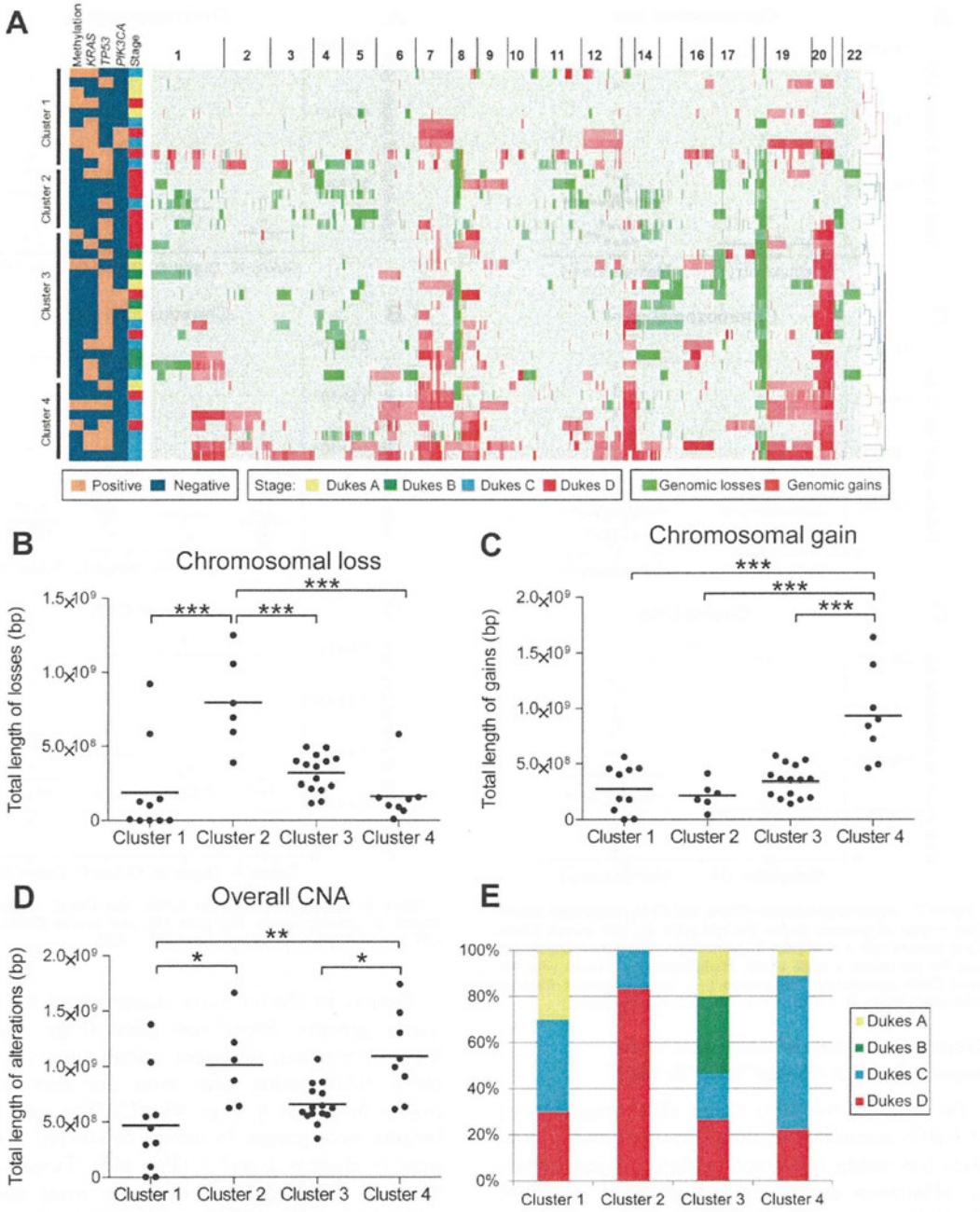

Figure 4. (A) Summarized results of unsupervised hierarchical analysis using the array-CGH data obtained from the 39 CRCs. Tumors were categorized into four clusters as indicated on the left. Tumors with frequent DNA methylation, *KRAS* mutation, *TP53* mutation, *PIK3CA* mutation, and Dukes' stages are also indicated on the left. (B–D) Total lengths of genomic losses and gains and overall CNAs in tumors within each cluster are shown. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. (E) Percentages of tumors at the respective Dukes' stages in each cluster are shown.

showed the greatest genomic gains. These results suggest that there is an inverse relationship between CIN and aberrant DNA methylation, and that there is a possible association between genomic losses and distant metastasis and between genomic gains and lymph node metastasis in CRC.

To determine whether a similar clustering pattern could be seen in an independent set of CRC samples, we carried out the same unsupervised

TABLE 2. Mean DNA Copy Number Alterations Within the Four Groups Obtained from the Unsupervised Clustering Analysis

Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	*P* value
Sample	10	6	15	8	
Age (mean ± SD)	66±14.7	69±11.2	69±10.1	63±15.2	
Sex					
Male	4	3	10	7	
Female	6	3	5	1	
Methylation	5/10*	0/6	2/15	2/8	
KRAS	6/10	1/6	5/15	3/8	
TP53	3/10	3/6	10/15	5/8	
PIK3CA	2/10	0/6	2/15	0/8	
18q loss	4/10*	6/6	15/15*	5/8	
20q gain	4/10*	3/6	14/15*	8/8	
18p loss	3/10*	4/6	13/15*	3/8	
8p loss	2/10*	4/6	11/15	4/8	
13q gain	1/10*	2/6	8/15	8/8*	
20p gain	2/10	0/6*	9/15	6/8*	
7p gain	4/10	1/6	5/15	6/8*	
7q gain	4/10	0/6*	7/15	5/8	
Loss (Mb)	189.4	795.8	318.5	162.6	<0.001
Gain (Mb)	277.4	216.3	339.1	930.3	0.001
Total CNA (Mb)	466.8	1012.1	657.6	1092.9	<0.001

Figures are the mean area of the CNAs per genome in each group, which were compared using one-way ANOVA. The total CNA is the sum of the losses and gains. Methylation, tumors with frequent DNA methylation.

*Significantly different from a random distribution, determined from the absolute value of the adjusted standardized residuals >1.96.

hierarchical clustering analysis using the publicly available data set (n = 121) reported by Nakao et al. (2004). We found that the CRC samples were subcategorized into four clusters, among which group 1 showed the highest levels of copy number gains, while group 3 showed the greatest genomic losses (Supporting Information Fig. 4).

DISCUSSION

CIN, which is inferred from DNA ploidy patterns, is the most common form of genomic instability in CRC and is defined by the presence of multiple structural or numerical chromosome changes (Rajagopalan and Lengauer, 2004; Grady and Carethers, 2008). Because mutations in mitotic checkpoint regulators are found in only a small population of CIN CRCs, the mechanism underlying CIN is largely unknown (Grady and Carethers, 2008). Moreover, no method for accurately evaluating the extent of CIN has yet been established. In this study, we used the total CNA length as a surrogate for CIN and analyzed its association with clinical and molecular variables. As described in an earlier study (Gaasenbeek et al., 2006), losses of heterozygosity (LOH) without accompanying copy number changes (e.g., somatic uniparental disomies) were not included as CNAs due to CGH array limitations. It is well established that estimates of CIN prevalence and mutation detection are influenced by the method used for analysis and the purity of the samples (Nakamura et al., 1994; Habano et al., 1996; Sugai et al., 2000; Cardoso et al., 2004; Issa, 2008). We therefore used a crypt isolation technique to avoid contamination by nontumorous cells. The high levels of marker gene methylation detected by quantitative pyrosequencing reflect the purity of the cancer cells in the isolated gland specimens and are indicative of the advantage of using the crypt isolation technique.

Several groups have reported an inverse relationship between CIMP and CIN (Goel et al., 2007; Cheng et al., 2008). Goel et al. (2007) found that CIN status measured as the LOH for eight microsatellite markers was inversely correlated with the methylation frequency at CIMP-related markers in sporadic CRCs, even those without MSI. In addition, Cheng et al. (2008) found that CIN measured as the number of chromosomal arms with gains/losses or LOH was inversely associated with CIMP status. In this study, we also observed an inverse relationship between concurrent methylation at multiple loci (which may represent CIMP) and CIN, although none of the tumors in this study exhibited *MLH1* methylation. Interestingly, we detected no significant difference in the magnitude of the chromosomal gains between tumors with frequent methylation and those without it. This suggests gains at certain genomic loci may be commonly involved in the pathogenesis of CRCs, irrespective of the methylation status. The presence of tumors with a high degree of chromosomal aberration in a subset of CIMP tumors may support this idea (Cheng et al., 2008).

There have been few studies in which the extent of chromosomal aberration during the progression of CRCs was analyzed as we have done in the present study. Hermsen et al. (2002) used conventional CGH to show that the progression of adenoma to carcinoma is associated with an increase in chromosomal aberration. A meta-analysis of 31 conventional CGH studies also showed that primary metastatic CRCs had significantly greater genomic alterations than non-metastatic CRCs (Diep et al., 2006). In addition, the results

of our array-CGH analysis suggest that accumulation of chromosomal losses may play an important role in the progression of CRCs from Dukes' A to Dukes' D. Unexpectedly, however, the overall level of CNAs was highest in Dukes' C tumors, with Dukes' D tumors showing somewhat less chromosomal alteration, which is probably attributable to the greater genomic gains in the Dukes' C group. Assuming that chromosomal gains are irreversible during CRC progression, at least a subset of Dukes' D tumors may have originated as Dukes' A or B tumors, not Dukes' C tumors with their larger chromosomal gains. Our results thus suggest that genomic losses may have a more significant impact on the metastatic properties of tumor cells than genomic gains, and that tumor cells that acquire losses at early stages in regions that are critical for metastasis may have a greater chance to metastasize to distant organs.

We also assessed the association between CNAs and tumor progression, including lymph node and distant metastasis. A previous meta-analysis of conventional CGH data suggests that losses at 4p are associated with the progression from Dukes' A tumors to Dukes' B–D tumors, and losses at 8p and gains at 7p and 17q correlate with liver metastasis (Diep et al., 2006). Our present findings that Dukes' D tumors frequently show losses at 4p and 8p and gains at 17q are consistent with those earlier findings. Importantly, our results seem to reflect the recent findings that losses at 4p and 8p are indicators of liver metastasis and a poor prognosis in CRC (Sheffer et al., 2009).

From an epigenetic viewpoint, Ju et al. (2011) reported that larger numbers of genes were methylated in stage I-III CRCs than in stage IV samples, and that CRCs at stages I-III exhibited methylation profiles that distinctly differed from the profiles of stage IV tumors. Together with the observation that even early stage tumors show intratumor heterogeneities, including ploidy pattern variation (Miyazaki et al., 1999), allelic imbalances (Boland et al., 1995; Sugai et al., 2005), and gene mutations (Baldus et al., 2010), tumor cells with metastatic potential may arise through early genetic and epigenetic lesions, as proposed in the "initiation" (Threadgill, 2005) or "parallel evolution" models (Gray, 2003).

In CRCs with distant metastasis, we frequently observed losses at 3q13.11, where *ACVR2B*, encoding activin A receptor type IIB, is located. Activins are growth and differentiation factors; they belong to the TGF-β superfamily and regulate cell differentiation, proliferation, and apoptosis in a variety of cancer cell types (Chen et al., 2006). Mutation of *ACVR2A*, which is 69% identical to *ACVR2B*, results in loss of its expression and is frequently found in MSI-positive CRCs (Jung et al., 2004). Furthermore, restoration of *ACVR2A* suppresses CRC cell growth, suggesting that it acts as a tumor suppressor (Jung et al., 2007). Together with an earlier report that expression of ACVR2B is weaker in CRC tissues than in normal colon mucosa (Babel et al., 2009), our findings suggest genomic loss of the *ACVR2B* locus may be involved in carcinogenesis and distant metastases in a subset of CRCs.

Consistent with the aforementioned meta-analysis of CGH data (Diep et al., 2006), we commonly observed genomic gains at 17q12, which encompasses a region encoding *ERBB2* and is frequently amplified and/or overexpressed in various types of cancer. One recent study showed that amplification of *ERBB2* leads to persistent activation of extracellular signal-regulated kinase 1/2 (ERK1/2) signaling, which in turn leads to resistance to treatment with cetuximab (Yonesaka et al., 2011). It has also been reported that the median overall survival was significantly longer for CRC patients without *ERBB2* amplification than for those with it (Yonesaka et al., 2011). Thus assessment of chromosomal aberrations at functionally important loci in CRCs could provide useful information for predicting distant metastasis and prognosis, as well as for the development of therapeutic strategies.

There have been a number of studies evaluating the utility of genomic instabilities, including CIN and CNAs, as prognostic markers in CRC (Pritchard and Grady, 2011). A recent meta-analysis confirmed that CIN (measured flow cytometrically as the presence of aneuploidy/polyploidy) is associated with an unfavorable prognosis (Walther et al., 2008). Our array-CGH analysis revealed an increase in copy number losses during the progression of tumor stage, suggesting that the total magnitude of genomic losses could potentially serve as a surrogate marker for genomic instabilities and a prognostic marker in CRC. Recently, Poulogiannis et al. (2010) carried out a hierarchical clustering analysis of array-CGH data obtained from 109 primary CRCs and showed that tumors could be categorized into four groups, in which tumors in groups I and II exhibited CNAs only infrequently, while those in groups III and IV exhibited an abundance of CNAs. CRCs in group I were characterized by a lack of CNAs and frequent MSI-positivity. This appears to be consistent with cluster 1 in our study, which was

enriched in tumors with frequent DNA methylation, and may suggest that the copy number profiles of CRCs with MSI are similar to those with frequent methylation but without *BRAF* mutation. They also reported that CRCs in group II, in which CNAs were somewhat more prevalent than in group I, were associated with a lack of lymph-node metastasis and a better prognosis, irrespective of MSI status (Poulogiannis et al., 2010). Thus group II may correspond to cluster 3 in our study, in which Dukes' B (lymph-node negative) tumors were significantly enriched. By contrast, most patients in our cluster 2 were staged as Dukes' D, with significantly greater genomic losses than tumors in other clusters. Furthermore, Dukes' C tumors were enriched in cluster 4, within which tumors exhibited the greatest genomic gains. Although the proportions of Dukes' C and D tumors in the study by Poulogiannis et al. differ from those in our study, the highest number of Dukes' D tumors were seen in group III and were characterized by significant genomic losses, while the majority of the Dukes' C tumors were in group IV and exhibited significant genomic gains. Although this distribution did not reach statistical significance, the tendency is consistent with our results. These findings suggest that cancers with gains are likely to metastasize to lymph nodes or unlikely to metastasize to distant sites. Thus a tumor's CNA pattern may define its metastatic behavior.

This study has several limitations, including a small sample size, the absence of tumors with *BRAF* mutation and/or *MLH1* methylation, and a lack of survival information; nonetheless, our results indicate several important findings. First, there is an inverse relationship between methylation status and the extent of the CNAs, particularly chromosomal losses. Second, tumor progression from Dukes' A to Dukes' C may be associated with the accumulation of CNAs, whereas a subset of Dukes' D tumors with smaller chromosomal gains may have been derived from Dukes' A or B tumors, but not Dukes' C tumors, which show significant genomic gains. Third, the different CNA patterns revealed by our hierarchical clustering analysis may be associated with different tumor behaviors, including local and distant metastasis, which suggests the presence of distinct molecular pathways in the development of CRC. Further study to clarify the differences between these subclasses will likely provide new insight into the molecular mechanisms that determine prognosis and the responses of CRCs to therapy.

ACKNOWLEDGMENTS

The authors thank Dr. William F. Goldman for editing the manuscript.

REFERENCES

Arai T, Kino I. 1989. Morphometrical and cell kinetic studies of normal human colorectal mucosa. Comparison between the proximal and the distal large intestine. Acta Pathol Jpn 39:725–730.

Babel I, Barderas R, Díaz-Uriarte R, Martínez-Torrecuadrada JL, Sánchez-Carbayo M, Casal JI. 2009. Identification of tumor-associated autoantigens for the diagnosis of colorectal cancer in serum using high density protein microarrays. Mol Cell Proteomics 8:2382–2395.

Baldus SE, Schaefer KL, Engers R, Hartleb D, Stoecklein NH, Gabbert HE. 2010. Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases. Clin Cancer Res 16:790–799.

Boland CR, Sato J, Appelman HD, Bresalier RS, Feinberg AP. 1995. Microallelotyping defines the sequence and tempo of allelic losses at tumour suppressor gene loci during colorectal cancer progression. Nat Med 1:902–909.

Cardoso J, Molenaar L, de Menezes RX, Rosenberg C, Morreau H, Möslein G, Fodde R, Boer JM. 2004. Genomic profiling by DNA amplification of laser capture microdissected tissues and array CGH. Nucleic Acids Res 32:e146

Chen YG, Wang Q, Lin SL, Chang CD, Chuang J, Ying SY. 2006. Activin signaling and its role in regulation of cell proliferation, apoptosis, and carcinogenesis. Exp Biol Med 231:534–544.

Cheng YW, Pincas H, Bacolod MD, Schemmann G, Giardina SF, Huang J, Barral S, Idrees K, Khan SA, Zeng Z, Rosenberg S, Notterman DA, Ott J, Paty P, Barany F. 2008. CpG island methylator phenotype associates with low-degree chromosomal abnormalities in colorectal cancer. Clin Cancer Res 14: 6005–6013.

Diep CB, Kleivi K, Ribeiro FR, Teixeira MR, Lindgj?rde OC, Lothe RA. 2006. The order of genetic events associated with colorectal cancer progression inferred from meta-analysis of copy number changes. Genes Chromosomes Cancer 45:31–41.

Dix B, Robbins P, Carrello S, House A, Iacopetta B. 1994. Comparison of p53 gene mutation and protein overexpression in colorectal carcinomas. Br J Cancer 70:585–590.

Douglas EJ, Fiegler H, Rowan A, Halford S, Bicknell DC, Bodmer W, Tomlinson IP, Carter NP. 2004. Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas. Cancer Res 64:4817–4825.

Fearon ER, Vogelstein B. 1990. A genetic model for colorectal tumorigenesis. Cell 61:759–767.

Gaasenbeek M, Howarth K, Rowan AJ, Gorman PA, Jones A, Chaplin T, Liu Y, Bicknell D, Davison EJ, Fiegler H, Carter NP, Roylance RR, Tomlinson IP. 2006. Combined array-comparative genomic hybridization and single-nucleotide polymorphism-loss of heterozygosity analysis reveals complex changes and multiple forms of chromosomal instability in colorectal cancers. Cancer Res 66:3471–3479.

Goel A, Nagasaka T, Arnold CN, Inoue T, Hamilton C, Niedzwiecki D, Compton C, Mayer RJ, Goldberg R, Bertagnolli MM, Boland CR. 2007. The CpG island methylator phenotype and chromosomal instability are inversely correlated in sporadic colorectal cancer. Gastroenterology 132:127–138.

Gray JW. 2003. Evidence emerges for early metastasis and parallel evolution of primary and metastatic tumors. Cancer Cell 4:4–6.

Grady WM, Carethers J. 2008. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology 135: 1079–1099.

Habano W, Sugai T, Nakamura S, Yoshida T. 1996. A novel method for gene analysis of colorectal carcinomas using a crypt isolation technique. Lab Invest 74:933–940.

Hermsen M, Postma C, Baak J, Weiss M, Rapallo A, Sciutto A, Roemen G, Arends JW, Williams R, Giaretti W, De Goeij A, Meijer G. 2002. Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. Gastroenterology 123:1109–1119.

Hinoue T, Weisenberger DJ, Lange CP, Shen H, Byun HM, Van Den Berg D, Malik S, Pan F, Noushmehr H, van Dijk CM, Tollenaar RA, Laird PW. 2012. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res 22:271–282.

Igarashi S, Suzuki H, Niinuma T, Shimizu H, Nojima M, Iwaki H, Nobuoka T, Nishida T, Miyazaki Y, Takamaru H, Yamamoto E, Yamamoto H, Tokino T, Hasegawa T, Hirata K, Imai K, Toyota M, Shinomura Y. 2010. A novel correlation between LINE-1 hypomethylation and the malignancy of gastrointestinal stromal tumors. Clin Cancer Res 16:5114–5123.

Issa JP. 2008. Colon cancer: It's CIN or CIMP. Clin Cancer Res 14:5939–5940.

Japanese Society for Cancer of the Colon and Rectum. 1997. Japanese Classification of Colorectal Carcinoma. Tokyo: Kanehara, pp.30–63.

Jass JR. 2007. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 50:113–130.

Jhawer M, Goel S, Wilson AJ, Montagna C, Ling YH, Byun DS, Nasser S, Arango D, Shin J, Klampfer L, Augenlicht LH, Perez-Soler R, Mariadason JM. 2008. PIK3CA mutation/PTEN expression status predicts response of colon cancer cells to the epidermal growth factor receptor inhibitor cetuximab. Cancer Res 68:1953–1961.

Jones AM, Douglas EJ, Halford SE, Fiegler H, Gorman PA, Roylance RR, Carter NP, Tomlinson IP. 2005. Array-CGH analysis of microsatellite-stable, near-diploid bowel cancers and comparison with other types of colorectal carcinoma. Oncogene 24:118–129.

Jover R, Nguyen TP, Pérez-Carbonell L, Zapater P, Payá A, Alenda C, Rojas E, Cubiella J, Balaguer F, Morillas JD, Clofent J, Bujanda L, Reñé JM, Bessa X, Xicola RM, Nicolás-Pérez D, Castells A, Andreu M, Llor X, Boland CR, Goel A. 2011. 5-Fluorouracil adjuvant chemotherapy does not increase survival in patients with CpG island methylator phenotype colorectal cancer. Gastroenterology 140:1174–1181.

Ju HX, An B, Okamoto Y, Shinjo K, Kanemitsu Y, Komori K, Hirai T, Shimizu Y, Sano T, Sawaki A, Tajika M, Yamao K, Fujii M, Murakami H, Osada H, Ito H, Takeuchi I, Sekido Y, Kondo Y. 2011. Distinct profiles of epigenetic evolution between colorectal cancers with and without metastasis. Am J Pathol 178:1835–1846.

Jung B, Doctolero RT, Tajima A, Nguyen AK, Keku T, Sandler RS, Carethers JM. 2004. Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers. Gastroenterology 126:654–659.

Jung BH, Beck SE, Cabral J, Chau E, Cabrera BL, Fiorino A, Smith EJ, Bocanegra M, Carethers JM. 2007. Activin type 2 receptor restoration in MSI-H colon cancer suppresses growth and enhances migration with activin. Gastroenterology 132: 633–644.

Meijer GA, Hermsen MA, Baak JP, van Diest PJ, Meuwissen SG, Beliën JA, Hoovers JM, Joenje H, Snijders PJ, Walboomers JM. 1998. Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation. J Clin Pathol 51:901–909.

Miyazaki M, Furuya T, Shiraki A, Sato T, Oga A, Sasaki K. 1999. The relationship of DNA ploidy to chromosomal instability in primary human colorectal cancers. Cancer Res 59: 5283–5285.

Nakamura S, Goto J, Kitayama M, Kino I. 1994. Application of the crypt-isolation technique to flow-cytometric analysis of DNA content in colorectal neoplasms. Gastroenterology 106:100–107.

Nakao K, Mehta KR, Fridlyand J, Moore DH, Jain AN, Lafuente A, Wiencke JW, Terdiman JP, Waldman FM. 2004. High-resolution analysis of DNA copy number alterations in colorectal cancer by array-based comparative genomic hybridization. Carcinogenesis 25:1345–1357.

Ogino S, Kawasaki T, Kirkner GJ, Loda M, Fuchs CS. 2006. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. J Mol Diagn 8:582–588.

Postma C, Koopman M, Buffart TE, Eijk PP, Carvalho B, Peters GJ, Ylstra B, van Krieken JH, Punt CJ, Meijer GA. 2009. DNA copy number profiles of primary tumors as predictors of response to chemotherapy in advanced colorectal cancer. Ann Oncol 20:1048–1056.

Poulogiannis G, Ichimura K, Hamoudi RA, Luo F, Leung SY, Yuen ST, Harrison DJ, Wyllie AH, Arends MJ. 2010. Prognostic relevance of DNA copy number changes in colorectal cancer. J Pathol 220:338–347.

Pritchard CC, Grady WM. 2011. Colorectal cancer molecular biology moves into clinical practice. Gut 60:116–129.

Rajagopalan H, Lengauer C. 2004. Aneuploidy and cancer. Nature 432:338–341.

Ried T, Knutzen R, Steinbeck R, Blegen H, Schröck E, Heselmeyer K, du Manoir S, Auer G. 1996. Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. Genes Chromosomes Cancer 15:234–245.

Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, Yan H, Gazdar A, Powell SM, Riggins GJ, Willson JK, Markowitz S, Kinzler KW, Vogelstein B, Velculescu VE. 2004. High frequency of mutations of the PIK3CA gene in human cancers. Science 304:554.

Sheffer M, Bacolod MD, Zuk O, Giardina SF, Pincas H, Barany F, Paty PB, Gerald WL, Notterman DA, Domany E.Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer.2009.Proc Natl Acad Sci USA 106:7131–7136.

Shen L, Catalano PJ, Benson AB III, O'Dwyer P, Hamilton SR, Issa JP. 2007a. Association between DNA methylation and shortened survival in patients with advanced colorectal cancer treated with 5-fluorouracil based chemotherapy. Clin Cancer Res 13:6093–6098.

Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, Hamilton SR, Issa JP. 2007b. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci USA 104:18654–18659.

Smith G, Carey FA, Beattie J, Wilkie MJ, Lightfoot TJ, Coxhead J, Garner RC, Steele RJ, Wolf CR. 2002. Mutations in APC, Kirsten-ras, and p53 - alternative genetic pathways to colorectal cancer. Proc Natl Acad Sci USA 99:9433–9438.

Sugai T, Habano W, Nakamura S, Uesugi N, Sasou S, Itoh C. 2000. A unique method for mutation analysis of tumor suppressor genes in colorectal carcinomas using a crypt isolation technique. Arch Pathol Lab Med 124:382–386.

Sugai T, Habano W, Jiao YF, Suzuki M, Takagi R, Otsuka K, Higuchi T, Nakamura S. 2005. Analysis of allelic imbalances at multiple cancer-related chromosomal loci and microsatellite instability within the same tumor using a single tumor gland from colorectal carcinomas. Int J Cancer 114:337–345.

Threadgill DW. 2005. Metastatic potential as a heritable trait. Nat Genet 37:1026–1027.

Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. 1999. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci USA 96:8681–8686.

Toyota M, Suzuki H, Sasaki Y, Maruyama R, Imai K, Shinomura Y, Tokino T. 2008. Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. Cancer Res 68:4123–4132.

Turnbull RB Jr, Kyle K, Watson FR, Spratt J. 1967. Cancer of the colon: the influence of the no-touch isolation technic on survival rates. Ann Surg 166:420–427.

Walther A, Houlston R, Tomlinson I. 2008. Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. Gut 57:941–950.

Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW. 2006. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 38:787–793.

Yonesaka K, Zejnullahu K, Okamoto I, Satoh T, Cappuzzo F, Souglakos J, Ercan D, Rogers A, Roncalli M, Takeda M, Fujisaka Y, Philips J, Shimizu T, Maenishi O, Cho Y, Sun J, Destro A, Taira K, Takeda K, Okabe T, Swanson J, Itoh H, Takada M, Lifshits E, Okuno K, Engelman JA, Shivdasani RA, Nishio K, Fukuoka M, Varella-Garcia M, Nakagawa K, Jänne PA. 2011. Activation of ERBB2 signaling causes resistance to the EGFR-directed therapeutic antibody cetuximab. Sci Transl Med 3:99ra86.

The American Journal of Pathology, Vol. 181, No. 5, November 2012

http://dx.doi.org/10.1016/j.ajpath.2012.08.007

Tumorigenesis and Neoplastic Progression

Molecular Dissection of Premalignant Colorectal Lesions Reveals Early Onset of the CpG Island Methylator Phenotype

Eiichiro Yamamoto,*† Hiromu Suzuki,*† Hiro-o Yamano,‡ Reo Maruyama,*† Masanori Nojima,§ Seiko Kamimae,† Takeshi Sawada,† Masami Ashida,† Kenjiro Yoshikawa,‡ Tomoaki Kimura,‡ Ryo Takagi,‡ Taku Harada,‡ Ryo Suzuki,* Akiko Sato,† Masahiro Kai,† Yasushi Sasaki,¶ Takashi Tokino,¶ Tamotsu Sugai,‖ Kohzoh Imai,** Yasuhisa Shinomura,* and Minoru Toyota†

From the First Department of Internal Medicine, and the Departments of Molecular Biology† and Public Health,§ and the Department of Medical Genome Science,¶ the Research Institute for Frontier Medicine, Sapporo Medical University, Sapporo; the Department of Gastroenterology,‡ Akita Red Cross Hospital, Akita; the Department of Pathology,‖ Iwate Medical University, Morioka; and The Advanced Clinical Research Center,** The Institute of Medical Science, The University of Tokyo, Tokyo, Japan*

The concept of the CpG island methylator phenotype (CIMP) in colorectal cancer (CRC) is widely accepted, although the timing of its occurrence and its interaction with other genetic defects are not fully understood. Our aim in this study was to unravel the molecular development of CIMP cancers by dissecting their genetic and epigenetic signatures in precancerous and malignant colorectal lesions. We characterized the methylation profile and *BRAF/KRAS* mutation status in 368 colorectal tissue samples, including precancerous and malignant lesions. In addition, genome-wide copy number aberrations, methylation profiles, and mutations of *BRAF*, *KRAS*, *TP53*, and *PIK3CA* pathway genes were examined in 84 colorectal lesions. Genome-wide methylation analysis of CpG islands and selected marker genes revealed that CRC precursor lesions are in three methylation subgroups: CIMP-high, CIMP-low, and CIMP-negative. Interestingly, a subset of CIMP-positive malignant lesions exhibited frequent copy number gains on chromosomes 7 and 19 and genetic defects in the *AKT/PIK3CA* pathway genes. Analysis of mixed lesions containing both precancerous and malignant components revealed that most aberrant methylation is acquired at the precursor stage, whereas copy number aberrations are acquired during the progression from precursor to malignant lesion. Our integrative genomic and epigenetic analysis suggests early onset of CIMP during CRC development and indicates a previously unknown CRC development pathway in which epigenetic instability associates with genomic alterations. *(Am J Pathol 2012, 181:1847–1861; http://dx.doi.org/10.1016/j.ajpath.2012.08.007)*

Colorectal cancer (CRC) is a leading cause of cancer mortality worldwide, but the incidence of CRC can be reduced through detection and removal of colorectal adenomas. Notably, however, most small colorectal polyps do not progress to malignancy; thus, the identification of precursor lesions likely to become cancerous is also extremely important for reducing CRC mortality.

CRCs are thought to arise, in part, through the accumulation of genetic changes, including mutations of oncogenes and tumor suppressor genes.[1] It is also generally accepted that CRCs can exhibit either of two genetic instabilities: chromosomal instability (CIN) or microsatellite instability (MSI).[2] In addition to these genetic changes, epigenetic alterations, including DNA methylation and histone

Supported by a Grant-in-Aid for the Third-Term Comprehensive 10-Year Strategy for Cancer Control (M.T. and H.S.), a Grant-in-Aid for Cancer Research from the Ministry of Health, Labor, and Welfare, Japan (M.T. and H.S.), the A3 foresight program from the Japan Society for Promotion of Science (H.S.), Grants-in-Aid for Scientific Research (A) from the Japan Society for Promotion of Science (K.I.), The Japanese Foundation for Research and Promotion of Endoscopy Grant (E.Y.), and a Research Grant from the Princess Takamatsu Cancer Research Fund (09-24119 to M.T.).

Accepted for publication August 1, 2012.

Supplemental material for this article can be found at http://ajp.amjpathol.org or at http://dx.doi.org/10.1016/j.ajpath.2012.08.007.

Address reprint requests to Hiromu Suzuki, M.D., Ph.D., Department of Molecular Biology, Sapporo Medical University, South-1, West-17, Chuo-ku, Sapporo 060-8556, Japan. E-mail: hsuzuki@sapmed.ac.jp.

Table 1. Primer Sequences Used in this Study

Gene	Primer/Target	Forward	Reverse	Product size (bp)
Methylation Analysis				
CDKN2A	Pyroseq PCR	5′-GGTTGTTTTGGTTGGTGTTTT-3′	5′-Bio-ACCCTATCCCTCAAATCCTCTAAAA-3′	169
	Sequence primer	5′-TTTTTTTGTTTGGAAAGAT-3′		
	Target	5′-ATYGYG-3′		
DFNA5	Pyroseq PCR	5′-GGYGGAGAGAGGGTTYGTT-3′	5′-Bio-RAACCCCTCCCRCAACCT-3′	91
	Sequence primer	5′-YGGGYGTTTTAGAGT-3′		
	Target	5′-YGYGGGATTGGTYGTYG-3′		
DKK2	Pyroseq PCR	5′-GGGTTTTTTGATTAATTAAGAGGAGA-3′	5′-Bio-TCTACAATAACTAAAAACAATCAAATAC-3′	179
	Sequence primer	5′-TAATTAAGAGGAGAGTTAAA-3′		
	Target	5′-TYGTYGAGATTTYGGYG-3′		
DLX4	Pyroseq PCR	5′-GGTTTYGGTTTAGTTTTGGATTTAGTT-3′	5′-Bio-CAATTCTACTCCCAAAAAACTCCCA-3′	182
	Sequence primer	5′-TGTTTYGTTTTATTTTAAGT-3′		
	Target	5′-TGGYGTTATYGTTYG-3′		
FZD10	Pyroseq PCR	5′-GGGATTTATTATAAAAGGAAGAGAAGAT-3′	5′-Bio-AATAATCCCCRACACCCCRAAAAC-3′	129
	Sequence primer	5′-AAAGGAAGAGAAGATGTATT-3′		
	Target	5′-TYGYG-3′		
GALNT14	Pyroseq PCR	5′-GAGYGGGAAAGTTTTTTTAGGTATAG-3′	5′-Bio-CCTAAACRCAACTCCCAAACCATC-3′	153
	Sequence primer	5′-GAAAGTTTTTTTAGGTATAG-3′		
	Target	5′-YGTYGTTTGGYG-3′		
IGF2BP1	Pyroseq PCR	5′-GAAGGGGTTTGTAGAGTTTTAGGGA-3′	5′-Bio-CCCACCCACCCTACAAAAAAAAACC-3′	157
	Sequence primer	5′-TTGAGTTTTTTATTTTTAGG-3′		
	Target	5′-YGGGAGATTATYG-3′		
IGFBP7	Pyroseq PCR	5′-AGGGTTYGGGGTAGGGGATTGGGGAT-3′	5′-Bio-AAAACCACACCCCRAAACRATAAAAACAC-3′	208
	Sequence primer	5′-YGGGTGTTYGTTTATTTT-3′		
	Target	5′-TYGAYGTTAGTAGGAGYGYGYGYG-3′		
KCNV1	Pyroseq PCR	5′-TAAGGAGAGGTAATTTTTTYGGGAGTT-3′	5′-Bio-CGCTAAAAAACATCTCTAACCCAATC-3′	150
	Sequence primer	5′-GGAGTTYGGGGAATTT-3′		
	Target	5′-YGGTYG-3′		
LRP1B	Pyroseq PCR	5′-GATGTAAGATTAGAYGTATTTTGTATTG-3′	5′-Bio-AACCAATCAACCTTCTCCTACCTAA-3′	148
	Sequence primer	5′-TATTGAAAAGTTAAGATATA-3′		
	Target	5′-YGGGYGTTTYGTTYGYG-3′		
MEOX2	Pyroseq PCR	5′-TAGAGTTTGGAGGGTAGAGTTGTTGT-3′	5′-Bio-ATTCCACTTCCTATCTCCTACTAAAC-3′	137
	Sequence primer	5′-GGGTAGAGTTGTTGTTTTT-3′		
	Target	5′-TYGGGYG-3′		
MINT1	Pyroseq PCR	5′-GGTTTTTTGTTAGYGTTTGTATTT-3′	5′-Bio-ATTAATCCCTCTCCCCTCTAAACTT-3′	133
	Sequence primer	5′-TTTAGTAAAAATTTTTTGGG-3′		
	Target	5′-GYGTTTGTTGTG-3′		
MINT2	Pyroseq PCR	5′-YGTTATGATTTTTTTGTTTAGTTAAT-3′	5′-Bio-TACACCAACTACCCAACTACCTC-3′	203
	Sequence primer	5′-TTTTGTTTAGTTAATTGAATTT-3′		
	Target	5′-GTYGTYGTTTYGAGTTTTAGG-3′		
MINT12	Pyroseq PCR	5′-YGGGTTATGTTTTATTTTTTGTGTTT-3′	5′-Bio-CTCAAAAAAATCAAACAACCAACCAA-3′	190

(*table continues*)

Y equals C or T, and R equals A or G.
Pyroseq, pyrosequencing.

Table 1. *Continued*

Gene	Primer/Target	Forward	Reverse	Product size (bp)
	Sequence primer	5′-TAATTYGGATTTTAAATTAAATA-3′		
	Target	5′-AAAYGTTTTTATTTT-3′		
MINT31	Pyroseq PCR	5′-GAYGGYGTAGTAGTTATTTTGTT-3′	5′-Bio-CATCACCACCCCTCACTTTAC-3′	184
	Sequence primer	5′-TGTAGTTTTAGGAGAGTGAATA-3′		
	Target	5′-AYGTTTAGGGGTGATGGTTTTAGTAAA-3′		
MIR34B	Pyroseq PCR	5′-GGTYGAGTGATTGTGGYGGGGG-3′	5′-Bio-CCTCCATCTTCTAAACRTCTCCCTTA-3′	176
	Sequence primer	5′-TAATYGTTTTTGGAATTT-3′		
	Target	5′-YGYGGGTYGAGGGGYGGGGYGGGYGYG-3′		
MLH1	Pyroseq PCR	5′-TTGGTATTTAAGTTGTTTAATTAATAGTTG-3′	5′-Bio-AAAATACCTTCAACCAATCACCTC-3′	119
	Sequence primer	5′-AGTTATAGTTGAAGGAAGAA-3′		
	Target	5′-YGTGAGTAYG-3′		
RASSF2	Pyroseq PCR	5′-GGTAGGGGTTGAAAAAGGTTAA-3′	5′-Bio-CRCRACTAAAAAACTACTTCAACT-3′	177
	Sequence primer	5′-GGYGTTYGGTTTTTA-3′		
	Target	5′-GTYGYGYGGTTATYG-3′		
RASSF5	Pyroseq PCR	5′-TYGTTATTAGTYGGGTATGGTTATGG-3′	5′-Bio-CRAAACCRCTCAAACTCTATAAATAAC-3′	110
	Sequence primer	5′-TATTYGTTATTATTGGATTT-3′		
	Target	5′-YGAGTYGTYGYG-3′		
SFRP1	Pyroseq PCR	5′-GTTTTGTTTTTTAAGGGGTGTTGAG-3′	5′-Bio-CTCCRAAAACTACAAAACTAAAATAC-3′	202
	Sequence primer	5′-GYGTTTGGTTTTAGTAAAT-3′		
	Target	5′-TTGYGYGGGGYGGTTTYGAGGGTTYG-3′		
SFRP2	Pyroseq PCR	5′-AATTTYGGATTGGGGTAAAATAAGTT-3′	5′-Bio-TTAAACAACAAACAAAAAAACCTAACC-3′	182
	Sequence primer	5′-YGTTTTYGTTAGTATTTGG-3′		
	Target	5′-TYGYGAGGTYGTTYGYG-3′		
SOX5	Pyroseq PCR	5′-GATTTGGAGGGAGYGGGAGTTTT-3′	5′-Bio-CAAAAACAAACAACACAATACRAATACA-3′	184
	Sequence primer	5′-GTYGTATTTTTYGGGG-3′		
	Target	5′-YGGGYGTYG-3′		
WIF1	Pyroseq PCR	5′-GTTTTYGTAGGTTTTTTGGTATTTAGG-3′	5′-Bio-GAACCATACTACTCAAAACCTCCTC-3′	174
	Sequence primer	5′-AGGTTTTTTGGTATTTAGG-3′		
	Target	5′-TYGGGAGGYGAYGYGTTTAGTYGTTTAAAYG-3′		
WNT5A	Pyroseq PCR	5′-ATATTTGGGGTTGGAAAGTTTTAATTAT-3′	5′-Bio-AACCRACAACAAAAACAAAACCTAATC-3′	149
	Sequence primer	5′-GGTTGGAAAGTTTTAATTAT-3′		
	Target	5′-YGTYGTYG-3′		
ZNF569	Pyroseq PCR	5′-TAGTYGATTGTAAGAAGGAAGTGTTT-3′	5′-Bio-CRCAAAAAAACTCAACCTAAATTTTACA-3′	199
	Sequence primer	5′-GGTTTTTGGGAAATGTA-3′		
	Target	5′-GTTYGGYG-3′		
Mutation Analysis				
AKT1	Pyroseq PCR	5′-Bio-AGTGTGCGTGGCTCTCACC-3′	5′-CATTCTTGAGGAGGAAGTAGCG-3′	83
	Sequence primer	5′-GCCAGGTCTTGATGTACT-3′		
	Target	5′-YCCCTA-3′		

modification, play critical roles in the development of CRC; in addition, an increasing number of genes involved in cell cycle control, DNA repair, tumor invasiveness, and the response to growth factors have been identified as targets of hypermethylation in CRC.[3–6] These epigenetic alterations are thought to be the main driving force in a subset of CRCs exhibiting concurrent hypermethylation of multiple loci, which is termed the CpG island methylator phenotype (CIMP).[7] CIMP-positive CRCs show characteristic clinicopathological and molecular features, including proximal tumor location, female sex, older age, high tumor grade, wild-type *TP53*, frequent *BRAF* and *KRAS* mutations, and MSI. In addition, several studies support the hypothesis that CRCs can be categorized into three subclasses based on aberrant CpG island methylation: CIMP-high (CIMP-H; also known as CIMP1), CIMP-low (CIMP-L; also known as CIMP2), and CIMP-negative (CIMP-N). CIMP-H CRCs are significantly associated with a *BRAF* mutation, *MLH1* methylation, and subsequent MSI.[7–11] A link between CIMP-L CRCs and *KRAS* mutations was first reported by Ogino et al,[12] and it was subsequently confirmed by other groups,[10,13] but much remains unknown about their respective molecular and clinicopathological features.

Reducing the incidence of cancers, such as CRC, will require a better understanding of the mechanisms underlying carcinogenesis and the molecular alterations occurring in premalignant lesions. For example, although experimental evidence has confirmed the presence of CIMP in CRCs, the role of CIMP in the progression of precancerous lesions toward cancer is not yet fully understood. In recent years, sessile serrated adenomas (SSAs) have been the origin of MSI-positive/CIMP-H cancers, which account for approximately 10% to 15% of sporadic CRCs.[14,15] In addition, studies have also shown that CIMP is frequently observed among MSI-negative CRCs.[10,16] From these and other studies of many tumors, a model was suggested in which CIN and epigenetic instability (CIMP) represent the two major pathways of CRC development, with up to 50% of CRCs being characterized as CIMP.[16,17] This means that the origin of a large fraction of CIMP cancers remains unclear.

High-resolution magnifying colonoscopy is a powerful diagnostic tool for detecting premalignant lesions. According to Kudo's classification, the pit patterns of non-neoplastic lesions are classified as type I (normal colon) or type II [hyperplastic polyp (HP)], whereas the pit patterns of neoplastic lesions are classified as types III, IV, and V.[18,19] Recently, we performed an integrative analysis of the morphological, pathological, and molecular signatures in colorectal precancerous lesions and identified a novel pit pattern (type II, open pits) that was specific to SSAs.[20] Those results depict an important relationship between morphological characteristics and molecular alterations that will significantly improve our ability to detect premalignant lesions. In the present study, our aim was to uncover the molecular evolution of CIMP cancers through an integrative analysis of many precursor and malignant colorectal lesions. Based on their genetic and epigenetic signatures, we propose a model in which CRCs develop via four distinct pathways. We also provide evidence that CIMP and CIN are not completely mutually exclusive, so that chromosomal aberrations may play important roles in a subset of CIMP cancers. These findings will improve our understanding of the pathogenesis of CRCs and could potentially contribute to better clinical management of premalignant lesions.

Materials and Methods

Study Population and Tissue Specimens

Colorectal tumor tissues were collected from Japanese patients who underwent endoscopic or surgical resection of a colorectal tumor at Akita Red Cross Hospital (Akita, Japan). A total of 368 specimens from 192 precursor lesions, 38 noninvasive carcinomas [carcinoma *in situ* (CIS)], 100 CRCs, and 38 samples of adjacent normal tissue were analyzed in this study. Informed consent was obtained from all patients before collection of the specimens. Approval of this study was obtained from the Institutional Review Board of Akita Red Cross Hospital and Sapporo Medical University (Sapporo, Japan). By using the standard phenol/chloroform procedure, genomic DNA was extracted from biopsy specimens obtained before endoscopic or surgical resection. CRC cell lines were maintained and treated with 5-aza-2′-deoxycytidine, as previously described.[21]

Endoscopic and Histological Analysis

High-resolution magnifying endoscopes (CF260AZI; Olympus, Tokyo, Japan) were used for all colonoscopic analyses. Colorectal subsite locations were defined as right side colon proximal to the splenic flexure (cecum, ascending colon, hepatic flexure, and transverse colon), left side colon distal to the splenic flexure (splenic flexure, descending colon, and sigmoid colon), and rectum (rectosigmoid and rectum). All detected colorectal tumors were observed at high magnification after staining with indigo carmine dye and 0.05% crystal violet. Surface microstructures were classified according to Kudo's pit pattern classification system.[18,19] Most often, one biopsy specimen was collected from each lesion for the extraction of genomic DNA. However, when two or more pit patterns were found in a single lesion (eg, adenoma to carcinoma transition), biopsy specimens were obtained for each respective pit pattern (see Supplemental Figure S1 at *http://ajp.amjpathol.org*). Thereafter, the lesions underwent endoscopic mucosal resection, endoscopic submucosal dissection, or surgical resection, after which histological analyses were performed (see Supplemental Figure S1 at *http://ajp.amjpathol.org*). Conventional adenomas, such as tubular adenoma and tubulovillous adenoma, were diagnosed using standard criteria. Serrated lesions, including HP, traditional serrated adenoma (TSA), and SSA, were classified based on criteria previously described by Torlakovic et al.[22] Serrated lesions that did not satisfy the criteria for SSA or TSA were defined as HP. Tumors were classified into three categories: precursor lesions (HP, tubular adenoma, tubulovillous adenoma, TSA, and SSA), CIS, and CRCs.

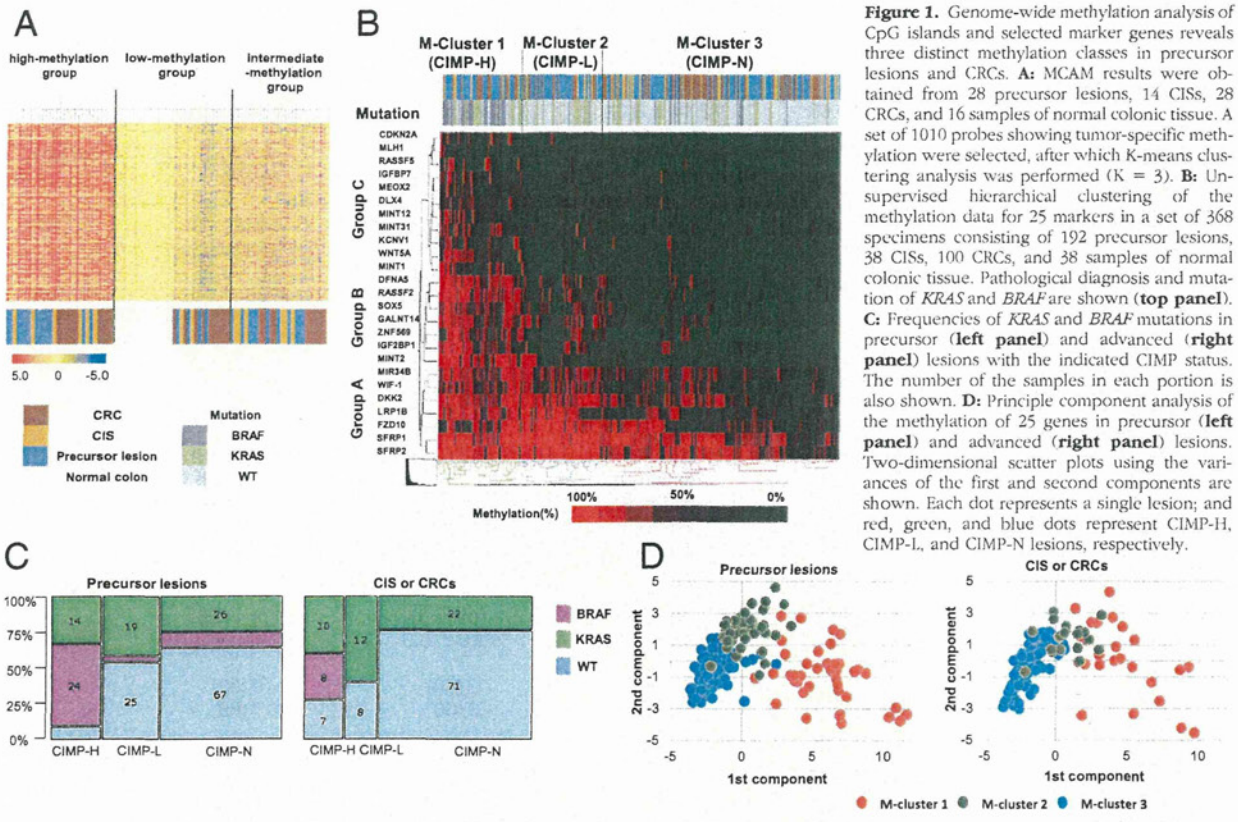

Figure 1. Genome-wide methylation analysis of CpG islands and selected marker genes reveals three distinct methylation classes in precursor lesions and CRCs. **A:** MCAM results were obtained from 28 precursor lesions, 14 CISs, 28 CRCs, and 16 samples of normal colonic tissue. A set of 1010 probes showing tumor-specific methylation were selected, after which K-means clustering analysis was performed (K = 3). **B:** Unsupervised hierarchical clustering of the methylation data for 25 markers in a set of 368 specimens consisting of 192 precursor lesions, 38 CISs, 100 CRCs, and 38 samples of normal colonic tissue. Pathological diagnosis and mutation of *KRAS* and *BRAF* are shown (**top panel**). **C:** Frequencies of *KRAS* and *BRAF* mutations in precursor (**left panel**) and advanced (**right panel**) lesions with the indicated CIMP status. The number of the samples in each portion is also shown. **D:** Principle component analysis of the methylation of 25 genes in precursor (**left panel**) and advanced (**right panel**) lesions. Two-dimensional scatter plots using the variances of the first and second components are shown. Each dot represents a single lesion; and red, green, and blue dots represent CIMP-H, CIMP-L, and CIMP-N lesions, respectively.

MCAM Data

Methylated CpG island amplification microarray (MCAM) analysis was performed as previously described.[23] A BioPrime Plus Array CGH Genomic Labeling System (Life Technologies, Carlsbad, CA) was used to label MCA amplicons from tumor samples with Alexa Fluor 647 (Life Technologies, Carlsbad, CA), and those from a pooled mixture of normal colonic tissue were labeled with Alexa Fluor 555 (Life Technologies). Labeled MCA amplicons were then hybridized to a custom human CpG island microarray (G4497A; Agilent Technologies, Santa Clara, CA), which included 15,134 probes covering 6157 unique genes. After washing, the array was scanned using an Agilent DNA Microarray Scanner (Agilent Technologies), and the data were processed using Feature Extraction software version 10.7 (Agilent Technologies), and analyzed using GeneSpring GX version 11 (Agilent Technologies). Unsupervised hierarchical clustering and statistical analyses were then performed.

Methylation Analysis by Bisulfite Pyrosequencing

Bisulfite pyrosequencing was performed as previously described.[24] Briefly, genomic DNA (1 μg) was modified with sodium bisulfite using an EpiTect Bisulfite Kit (Qiagen, Hilden, Germany). Pyrosequencing was then performed using a PSQ96 system with a PyroGold reagent Kit (Qiagen), after which the results were analyzed using Q-CpG software version 1.0.9 (Qiagen). Unsupervised hierarchical clustering, principal component analysis, and correspondence analysis of validation data were performed using JMP version 8 (SAS Institute Inc., Cary, NC). For the statistical analysis, quantitative methylation levels of each gene were converted to *z* scores, which were defined as follows: (Methylation Level in Each Sample - Mean Methylation Level in All Samples)/(SD of Methylation Levels in All Samples). Primer sequences are listed in Table 1.

Mutation and MSI Analysis

Mutations in codon 600 of *BRAF* and codons 12 and 13 of *KRAS* were examined by pyrosequencing using *BRAF* and *KRAS* pyrokits (Qiagen), respectively, according to the manufacturer's instructions. Mutation of *PIK3CA*, *AKT2*, and *PDK1* was analyzed by direct sequencing, as previously described.[25] Mutation of *TP53* was initially detected by PCR–single-stranded conformational polymorphism analysis, followed by direct sequencing, as previously described.[26] Mutation of *AKT1* was analyzed by pyrosequencing using the primer sequences listed in Table 1. MSI was assessed as previously described,[27] using primers proposed by the National Cancer Institute Workshop on Microsatellite Instability (*BAT25*, *BAT26*, *D5S346*, *D2S123*, and *D17S250*).[28] MSI was defined by the presence, in the tumor sample, of bands that were abnormal in size, compared with a corresponding normal

Table 2. Clinicopathological Features of the Colorectal Tumors Used in this Study

Feature	Total (N = 192)	CIMP-H (n = 42)	CIMP-L (n = 46)	CIMP-N (n = 104)	P value
Precursor Lesion					
Age (years)	69.61 ± 9.68	72.14 ± 8.01	72.43 ± 9.82	67.34 ± 9.73	<0.05
Sex					
F	73 (38.02)	21 (50)	26 (56.52)	26 (25)	<0.001
M	119 (61.98)	21 (50)	20 (43.48)	78 (75)	
Location					
Right	92 (47.92)	32 (76.19)	27 (58.7)	33 (31.73)	<0.001
Left	40 (20.83)	5 (11.9)	7 (15.22)	28 (26.92)	
Rectum	60 (31.25)	5 (11.9)	12 (26.09)	43 (41.35)	
Size (mm)	12.98 ± 9.61	13.73 ± 8.44	16.67 ± 12.55	8.38 ± 6.48	<0.001
Morphological characteristics					
Protruding type	82 (42.71)	17 (40.48)	21 (45.65)	44 (42.31)	
Flat type	110 (57.29)	25 (59.52)	25 (54.35)	60 (57.69)	
Depressed type	0	0	0	0	
Histological features					
HP	28 (14.58)	2 (4.76)	5 (10.87)	21 (20.19)	<0.001
SSA	29 (15.1)	26 (61.90)	1 (2.17)	2 (1.92)	
TSA	25 (13.02)	6 (14.29)	5 (10.87)	14 (13.46)	
Tubular adenoma	53 (17.6)	2 (4.76)	7 (15.22)	44 (42.31)	
Tubulovillous adenoma	57 (29.69)	6 (14.29)	28 (60.87)	23 (22.12)	
CIS					
	(N = 38)	(n = 8)	(n = 5)	(n = 25)	
Age (years)	66.71 ± 13.77	75.75 ± 7.78	73.6 ± 5.5	62.44 ± 14.61	<0.05
Sex					
F	13 (34.21)	3 (37.5)	2 (40)	8 (32)	
M	25 (65.79)	5 (62.5)	3 (60)	17 (68)	
Location					
Right	18 (47.37)	6 (75)	4 (80)	8 (32)	<0.05
Left	9 (23.68)	0	1 (20)	8 (32)	
Rectum	11 (28.95)	2 (25)	0	9 (36)	
Size (mm)	17.66 ± 9.75	20.63 ± 12.96	19.6 ± 10.21	16.32 ± 8.64	
Morphological features					
Protruding type	20 (52.63)	6 (75)	1 (20)	13 (52)	
Flat type	16 (42.11)	2 (25)	4 (80)	10 (40)	
Depressed type	2 (5.26)	0	0	2 (8)	
CRC					
	(N = 100)	(n = 17)	(n = 15)	(n = 68)	
Age (years)	67.81 ± 12.48	71 ± 9.54	72.13 ± 12.78	66.06 ± 12.82	
Sex					
F	43 (43)	14 (82.35)	8 (53.33)	21 (30.88)	<0.01
M	57 (57)	3 (17.65)	7 (46.67)	47 (69.12)	
Location					
Right	48 (48)	16 (94.12)	6 (40)	26 (38.24)	<0.001
Left	22 (22)	1 (5.88)	3 (20)	18 (26.47)	
Rectum	30 (30)	0 (0)	6 (40)	24 (35.29)	
Stage (UICC)					
I	38 (38)	8 (47.06)	3 (20)	27 (39.71)	
II	31 (31)	4 (23.53)	5 (33.33)	22 (32.35)	
III	23 (23)	5 (29.41)	7 (46.67)	11 (16.18)	
IV	8 (8)	0	0	8 (11.76)	

F, female; M, male; UICC, Union for International Cancer Control.

sample. A tumor sample was defined as MSI positive when two or more markers showed instability.

Array-Based Comparative Genomic Hybridization

Array-based comparative genome hybridization (array CGH) analysis was performed as previously described.[29] Briefly, 500 ng of genomic DNA from colorectal tumor specimens and sex-matched reference DNA from noncancerous colonic mucosa were labeled with Cy5 and Cy3, respectively, using a Genomic DNA Enzymatic Labeling Kit (Agilent Technologies), and were then hybridized to Human Genome CGH Microarray Kit 180A (G4449A; Agilent Technologies). DNA copy number aberrations (CNAs) were identified using the ADM-2 algorithm included in the Agilent Genomic Workbench software version 5 (Agilent Technologies). A copy number loss was defined as a log 2 ratio of <-0.5, and a copy number gain was defined as a log 2 ratio of >0.5. Unsupervised hierarchical clustering of CNAs was performed using JMP, version 8. The microarray data in this study

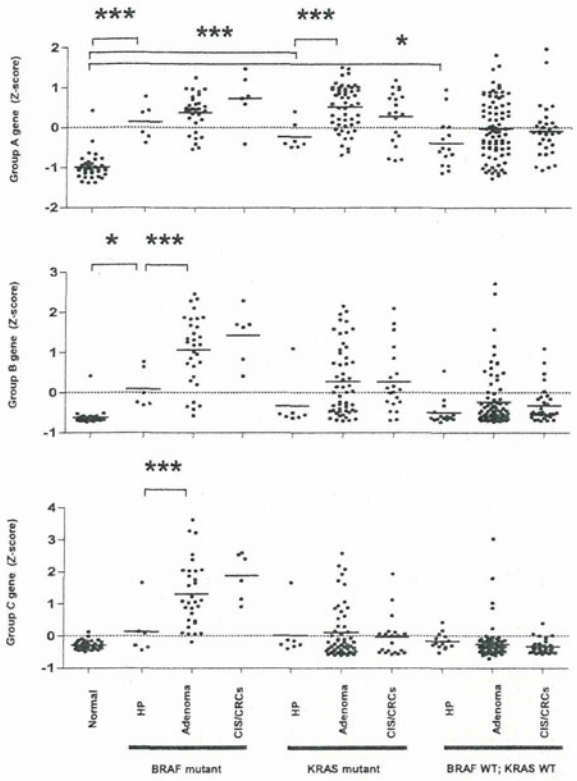

Figure 2. Association between methylation and *BRAF*/*KRAS* mutations. The levels of methylation of groups A to C genes in normal colon, precursor, and malignant lesions are shown as *z* scores. Colorectal lesions are divided into three groups, according to their *BRAF* or *KRAS* mutation status. Among the *BRAF* mutants, levels of methylation of group B and C genes are significantly higher in adenomas than HPs but are not further up-regulated in advanced lesions. $^{*}P < 0.05$, $^{***}P < 0.001$.

have been submitted to the Gene Expression Omnibus, and the accession number is GSE35534.

Statistical Analysis

To compare differences in continuous variables between groups, *t*-tests or an analysis of variance with a post hoc Tukey's test was performed. A Fisher's exact test or a χ^2 test was used for analysis of categorical data. $P < 0.05$ (two sided) was considered statistically significant. Statistical analyses were performed using JMP, version 8, and SPSS statistics 18 (IBM Corporation, Somers, NY).

Results

Three Methylation Subclasses in Precancerous and Malignant Colorectal Tumors

To clarify the epigenetic changes occurring early during colorectal tumorigenesis, we first performed global methylation analysis using MCAM in a series of normal colonic tissues ($n = 16$), precursor lesions (HP, $n = 3$; SSA, $n = 5$; TSA, $n = 3$; tubular adenoma, $n = 6$; tubulovillous adenoma, $n = 10$), CISs ($n = 14$), and CRCs ($n = 28$). Hierarchical clustering analysis using the MCAM results identified 1010 probe sets that detected frequent hypermethylation in tumor tissues (see Supplemental Figure S2 at *http://ajp.amjpathol.org*). Subsequent K-means clustering analysis using these probe sets revealed that the samples could be clearly categorized into three subclasses based on the level of their methylation (Figure 1A); subclasses with high and intermediate methylation were presumed to reflect CIMP-H and CIMP-L tumors, respectively. Among the 1010 probe sets, 538 unique genes were in the high-methylation group, whereas 259 genes were in the intermediate-methylation group (see Supplemental Figure S2C and Supplemental Tables S1 and S2 at *http://ajp.amjpathol.org*). A subset of the precursor lesions could also be categorized into these subclasses, indicating that CIMP-like methylation patterns were already established early during carcinogenesis.

To further characterize the genes that acquired methylation progressively during CIMP pathway tumorigenesis, we next performed MCAM analysis with a series of precursor lesions in which CIMP-N flat components were present, along with CIMP-positive protruding components within the same lesions (see Supplemental Figure S2D at *http://ajp.amjpathol.org*). Because both components were presumed to arise from the same origin, these lesions could represent an ideal model for analyzing the molecular progression of CIMP cancers. CIMP status was defined using classic CIMP markers (*MINT1*, *MINT2*, *MINT12*, *MINT31*, and *CDKN2A*), and tumors with methylation of three or more markers were defined as CIMP. When we analyzed three pairs of precursor lesions using MCAM, we identified 36 unique genes that were differentially methylated between CIMP-positive and CIMP-N precursor lesions (see Supplemental Table S3 at *http://ajp.amjpathol.org*). These genes were potentially the earliest targets of aberrant methylation during CIMP pathway tumorigenesis, and most of them overlapped with the genes identified in the initial MCAM analysis (see Supplemental Figure S2E at *http://ajp.amjpathol.org*).

Methylation Profiling Identified CIMP in Precancerous Lesions

Based on the results previously summarized, we selected a series of marker genes to characterize the methylation profile of precursor and malignant lesions. We initially selected 14 genes (*LRP1B*, *CDKN2A*, *WNT5A*, *MEOX2*, *ZNF569*, *GALNT14*, *SOX5*, *DFNA5*, *DLX4*, *SFRP2*, *WIF1*, *FZD10*, *KCNV1*, and *IGF2BP1*) identified in the MCAM analysis (see Supplemental Figure S2 at *http://ajp.amjpathol.org*). Among them, *IGF2BP1*, *KCNV1*, *DLX4*, *GALNT14*, and *ZNF569* were not previously reported to be methylated in CRCs, but RT-PCR analysis using multiple CRC cell lines confirmed that they were frequent targets of epigenetic silencing in CRC (see Supplemental Figure S3 at *http://ajp.amjpathol.org*). In addition, we selected 11 well-characterized markers (*MLH1*, *SFRP1*, *IGFBP7*, *DKK2*, *MIR34B*, *MINT1*, *MINT2*, *MINT12*, *MINT31*, *RASSF2*, and *RASSF5*) used for methylation analysis.[4,7,21,24,30–32]

Table 3. Clinicopathological Features of the Colorectal Tumors Analyzed Using Array CGH

Feature	C-cluster 1 (n = 49)	C-cluster 2 (n = 13)	C-cluster 3 (n = 22)	P value
Age (years)	69.51 ± 11.01	69.54 ± 8.97	67 ± 10.55	
Sex				
F	14 (28.57)	9 (69.23)	4 (18.18)	
M	35 (71.43)	4 (30.77)	18 (81.82)	
Location				
Right	27 (55.1)	9 (69.23)	6 (27.27)	
Left	8 (16.33)	4 (30.77)	7 (31.82)	
Rectum	14 (28.57)	0	9 (40.91)	
Precursor lesions				
HP	1 (2.04)	0	0	
SSA	2 (4.08)	0	0	
TSA	7 (14.29)	1 (7.69)	0	
Tubular adenoma	3 (6.12)	1 (7.69)	0	
Tubulovillous adenoma	1 (2.04)	0	1 (4.55)	
CIS	0	0	2 (9.09)	
CRCs	9 (18.37)	2 (15.38)	8 (36.36)	
SSA + CIS				
SSA portion	2 (4.08)	0	0	
CIS portion	1 (2.04)	1 (7.69)	0	
Tubulovillous adenoma + CIS				
Tubulovillous adenoma portion	12 (24.49)	1 (7.69)	0	
CIS portion	4 (8.16)	6 (46.15)	3 (13.64)	
Tubular adenoma + CIS				
Tubular adenoma portion	5 (10.2)	0	3 (13.64)	
CIS portion	2 (4.08)	1 (7.69)	5 (22.73)	
KRAS				
Mut	22 (45)	11 (85)	10 (45)	
Wt	27 (55)	2 (15)	12 (55)	
BRAF				
Mut	10 (20)	1 (8)	0	
Wt	39 (80)	12 (92)	22 (100)	
CIMP				
High	17 (35)	4 (31)	0	<0.01
Low	11 (22)	4 (31)	1 (5)	
Negative	21 (43)	5 (38)	21 (95)	

F, female; M, male; Mut, mutated; Wt, wild type.

We performed quantitative bisulfite pyrosequencing of the 25 markers in a total of 330 specimens consisting of 192 precursor lesions, 38 CISs, and 100 CRCs (Table 2). Consistent with the MCAM results, unsupervised hierarchical clustering using the pyrosequencing data revealed that, in addition to malignant lesions (CISs and CRCs), precursor lesions could also be divided into three relative subclasses (M-clusters 1 to 3; Figure 1B). The *BRAF* mutation was significantly enriched in M-cluster 1 tumors, in which most of the genes were highly methylated, suggesting that this subclass corresponded to CIMP-H (Figure 1B). M-cluster 2 included tumors with intermediate levels of methylation and a prevalent *KRAS* mutation, which corresponded to CIMP-L, whereas M-cluster 3 tumors exhibited the lowest methylation levels, which corresponded to CIMP-N. Among the precursor lesions, 37 (19.3%) exhibited a *BRAF* mutation, and most of those [24 (64.9%) of 37] were categorized as CIMP-H (Figure 1C). In CISs and CRCs, the *KRAS* mutation was most enriched in the CIMP-L group, but this tendency was less apparent among the precursor lesions (Figure 1C).

We next used principle component analysis to further evaluate our bisulfite pyrosequencing results and found that the first and second components accounted for 53.1% of the total variance (Figure 1D; see also Supplemental Table S4 at *http://ajp.amjpathol.org*). Two-dimensional plotting then showed that the characteristic pattern for each M-cluster was shared by the precursor and malignant lesions (Figure 1D; see also Supplemental Figure S4A at *http://ajp.amjpathol.org*).

Our hierarchical clustering analysis also showed that the marker genes could be categorized into three subgroups (Figure 1B). Group A genes, which were methylated in most of the samples, corresponded to the type A (age-related) genes originally proposed by Toyota et al.[7] Group B and C genes appeared to correspond to type C (cancer-specific) genes, whereas group C genes in this study were more strongly associated with CIMP-H. Interestingly, among *BRAF*-mutant precursor lesions, adenomas showed much higher levels of methylation of groups B and C genes than HPs (Figure 2). Similar results were obtained with *KRAS*-mutant precursors, although the difference was not statistically significant (Figure 2). By contrast, methylation was minimally up-regulated in adenomas in which both *BRAF* and *KRAS* were wild type, suggesting that accumulation of aberrant methylation, in concert with a *BRAF* or a *KRAS* mutation, may promote the progression from benign tumors to precancerous lesions.