Epigenetic Patterns in the Progression of Esophageal Adenocarcinoma


ABSTRACT

Esophageal adenocarcinoma (EAC) arises after normal squamous mucosa undergoes metaplasia to specialized columnar epithelium (intestinal metaplasia or Barrett’s esophagus), which can then ultimately progress to dysplasia and subsequent malignancy. Epigenetic studies of this model have thus far been limited to the DNA methylation analysis of a few genes. In this study, we analyzed a panel of 20 genes using a quantitative, high-throughput methylation assay, MethyLight. We used this broader approach to gain insight into concordant methylation behavior between genes and to generate epigenomic fingerprints for the different histological stages of EAC. Our study included a total of 104 tissue specimens from 51 patients with different stages of Barrett’s esophagus and/or associated adenocarcinoma. We screened 84 of these samples with the full panel of 20 genes and found distinct classes of methylation patterns in the different types of tissue. The most informative genes were those with an intermediate frequency of significant hypermethylation [ranging from 15% (CDKN2A) to 60% (MGMT) of the samples]. This group could be further subdivided into three classes, according to the absence (CDKN2A, ESRI, and MYOD1) or presence (CALCA, MGMT, and TIMP3) of methylation in normal esophageal mucosa and stomach, or the infrequent methylation of normal esophageal mucosa accompanied by methylation in all normal stomach samples (APC). The other genes were less informative, because the frequency of hypermethylation was below 5% (ARF, CDH1, CDKN2B, GSTP1, MLH1, PTGS2, and THBS1), completely absent (CTNNB1, RB1, TGFB2, and TYMS1), or ubiquitous (HIC1 and MTHFR), regardless of tissue type. Each class undergoes unique epigenetic changes at different steps of disease progression of EAC, suggesting a step-wise loss of multiple protective barriers against CpG island hypermethylation. The aberrant hypermethylation occurs at many different loci in the same tissues, suggestive of an overall deregulation of methylation control in EAC tumorigenesis. However, we did not find evidence for a distinct group of tumors with a CpG island methylator phenotype. Finally, we found that normal and metaplastic tissues from patients with evidence of associated dysplasia or cancer had a significantly higher incidence of hypermethylation than similar tissues from patients with no further progression of their disease. The fact that the samples from these two groups of patients were histologically indistinguishable, yet molecularly distinct, suggests that the occurrence of such hypermethylation may provide a clinical tool to identify patients with premalignant Barrett’s who are at risk for further progression.

INTRODUCTION

DNA methylation patterns are frequently altered in human cancers. These methylation changes include genome-wide hypomethylation as well as regional hypermethylation (1). Aberrant hypermethylation in cancer cells often occurs at CpG islands, which are generally protected from methylation in normal tissues. Methylation patterns of genes can provide different types of useful information about a cancer cell. First, each tumor type (i.e., breast, colon, esophagus, and so forth) has a characteristic set of genes with an increased propensity to become methylated (2). For example, RB1 is known to be hypermethylated in retinoblastoma (3, 4), but not in acute myelogenous leukemia (5, 6). Second, an individual tumor within a single patient has a unique epigenetic fingerprint reflective of the evolution of that tumor as compared with a tumor of the same type in a different patient (2). Determining tumor-type-specific and patient-specific fingerprints may not only shed light on coordinate patterns of CpG island methylation changes at multiple loci during different steps of a disease, but may also provide biomarkers that can be used diagnostically. For instance, such epigenome maps may be invaluable in cancer detection, in cancer chemoprediction and in prognostics (1, 7).

The incidence of EAC has increased rapidly in the Western World over the past three decades (8, 9). EAC arises from a multistep process whereby normal squamous mucosa undergoes metaplasia to specialized columnar epithelium (IM or Barrett’s esophagus), which then ultimately progresses to DYS and subsequent malignancy (10, 11). Epigenetic studies of this model have thus far been limited to the DNA methylation analysis of a few genes (12–14). Most studies of epigenetic alterations in cancer have focused primarily on either a very small set of known genes (1, 15) or on the global analysis of unknown CpG islands (2). We have previously used a targeted approach to show that hypermethylation of the genes APC, CDKN2A, and ESRI occur as early as the premalignant Barrett’s esophagus stage (14). However, this study was limited to six patients. The general frequency with which these methylation changes occurred could not be accurately resolved.

It has previously been reported that a subset of colorectal and gastric tumors display a CIMP, whereby aberrant hypermethylation changes are widespread, affecting multiple loci in a single tumor (16, 17). This is reflected in a bimodal distribution of the frequency of the number of genes methylated in a group of tumors (16). CIMP tumors are a distinct group of tumors that are defined by a high degree of concordant CpG island hypermethylation of genes exclusively methylated in cancer, or type C genes. CIMP is now thought to be a new, distinct, yet major pathway of tumorigenesis (16, 17). The role of the CIMP pathway in the tumor evolution of EAC is still unknown because the previous epigenetic studies analyzed only one (12, 13) or only a few genes (14).

In this report, we have now combined the advantages of both the targeted and the comprehensive approaches by analyzing 20 different genes (see Table 1) using a quantitative, high-throughput methylation assay, MethyLight (14, 18, 19): (a) to more extensively characterize the methylation changes in EAC; (b) to generate epigenomic fingerprints for the different histological stages of EAC; (c) to identify epigenetic biomarkers useful in disease diagnosis and prevention; and (d) to determine whether CIMP is a contributor to the tumorigenesis of EAC tumors. We screened 104 tissue specimens from 51 patients with different stages of Barrett’s esophagus and/or associated adeno-
carcinoma. We found that the 20 genes segregated into classes of similar epigenetic behavior. Each class undergoes unique epigenetic changes at different steps of disease progression of EAC suggesting a step-wise loss of multiple protective barriers against CpG island hypermethylation. These epigenetic profiles may prove to be useful as a clinical tool to monitor or diagnose Barrett’s esophagus or EAC. Furthermore, we found no clear evidence of a CIMP in the esophageal tumors analyzed.

MATERIALS AND METHODS

Sample Collection. Multiple tissue samples [NE, normal stomach (5), IM, DYS, and/or adenocarcinoma (7)] from a total of 51 patients (age range, 39–86 years), with either adenocarcinoma or IM as the most advanced stage of disease, were collected. The initial set of samples analyzed included biopsies from 31 patients that were collected fresh and subdivided such that a part of each specimen was immediately frozen in liquid nitrogen and a part also embedded in paraffin for histopathological examination by a pathologist (K. W.). Normal esophageal tissue was collected from every patient 10 cm or more away from the diseased areas. Frozen section examination of the frozen tissues was performed if the diagnosis was uncertain. TNM staging was used to classify the stage of each adenocarcinoma. A second set of samples were obtained for a follow-up study of 20 cases. Two groups of IM samples were collected: patients that had only IM as the most advanced stage of disease (8 patients) and patients that had IM with associated DYS/adenocarcinoma (12 patients). Five-µm H&E slides for each sample were prepared and examined by a pathologist (K. W.) to verify and localize the IM tissue. We excluded cases from this follow-up study that showed any signs of DYS or adenocarcinoma in the paraffin block used for analysis. The IM tissues were carefully microdissected away from other cell types from a 30-µm section adjacent to the 5-µm H&E section. All of the specimens were classified according to the highest grade histopathological lesion present in that sample. Approval for this study was obtained from the Institutional Review Board of the University of Southern California Keck School of Medicine.

Nucleic Acid Isolation. Genomic DNA was isolated from the frozen-tissue biopsies by a simplified protease K digestion method (20). The DNA from the paraffin tissues was extracted in lysis buffer (Tris-HCl (100 mM; pH 8), EDTA (10 mM), proteinase K (1 mg/ml) overnight at 50°C (21).

Sodium Bisulfite Conversion. Sodium bisulfite conversion of genomic DNA was performed as described previously (22). The beads were incubated for 14 h at 50°C to ensure complete conversion.

MethyLight Analysis. After sodium bisulfite conversion, the methylation analysis was performed by the fluorescence-based, real-time PCR assay MethyLight as described previously (14, 18, 19). Two sets of primers and probes, designed specifically for bisulfite-converted DNA, were used: a methylated set for the gene of interest and a reference set, β-actin (ACTB), to normalize for input DNA. Specificity of the reactions for methylated DNA were confirmed separately using human sperm DNA (with very low levels of CpG island methylation) and SsoI (New England Biolabs)-treated sperm DNA (heavily methylated) as described previously (14). The percentage of fully methylated molecules at a specific locus was calculated by dividing the GENE:ACTB ratio of a sample by the GENE:ACT ratio of SsoI-treated sperm DNA and multiplying by 100. We use the abbreviation PMR to indicate this measurement. The methylation analysis on the paraffin-microdissected samples was performed after bisulfite treatment as described above by an investigator (C. A. E.) blind to the associated DYS status of the samples. The primers and probes used in the methylation analysis are listed in Table 2.

Statistics. The PMR values obtained by MethyLight (see above) were dichotomized at 4 PMR for statistical purposes as described previously (14). Dichotomization moderates the quantitative impact of gene loci with different levels of hypermethylation, resulting in a more reliable cross-gene comparison of hypermethylation frequencies. A dichotomization point of 4 PMR was selected because it gave the best discrimination between normal and malignant tissues, across the board for all CpG islands (14). Samples containing 4 PMR or higher were designated as methylated and given a value of 1, whereas samples containing less than 4 PMR were designated as unmethylated and given a value of 0. The cumulative value of genes methylated in each class or for all of the 19 genes was then used as a continuous variable in a Fisher’s PLSD test, adapted for use with unequal sample sizes (SAS Statview software) to obtain P-values. The different parameters such as tissue type, presence of asso-
Three oligonucleotides are used in every reaction: two locus-specific PCR primers flanking an oligonucleotide probe with a 5' fluorescent reporter dye (6FAM) and a 3' quencher dye (TAMRA) (48). The Genbank accession number for each sequence is listed with the corresponding PCR amplicon location within that sequence. The percentage of guanine-cytosine (%GC) content, CpG observed/expected value, and CpG/GpC ratio of 200 bp encompassing the Methylight amplicon are indicated for each gene. The reaction type is designated M for methylation reaction and C for control reaction. The bisulfite-treated DNA strand [top (T) or bottom (B)] and amplicon orientation [parallel (P) or antiparallel (A)] is also indicated. All of the primer and probe sequences are listed in the 5' to 3' direction.

<table>
<thead>
<tr>
<th>HUGO gene</th>
<th>Genbank accession no.</th>
<th>Bisulfite (T or B strand)</th>
<th>Amplicon location (Genbank numbering)</th>
<th>Amplicon location relative to transcription start (bp)</th>
<th>% GC content</th>
<th>Obs/Exp</th>
<th>CpG</th>
<th>CpG/GpC</th>
<th>Forward primer sequence (5'→3')</th>
<th>Probe sequence (5'→3')</th>
<th>Reverse primer sequence (5'→3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDKN2Aa</td>
<td>NM_000077</td>
<td>T</td>
<td>66–113bp</td>
<td>9' +9' +69</td>
<td>72</td>
<td>1.68</td>
<td>0.81</td>
<td>6FAM-ACCCGACCCGACCGG-TAMRA</td>
<td>AACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESR1</td>
<td>X62462</td>
<td>T</td>
<td>2784–2884bp</td>
<td>+14/+114</td>
<td>69</td>
<td>1.52</td>
<td>0.83</td>
<td>6FAM-CGTGCTTGGTGGATG-TGTT</td>
<td>GTGAGTTGGATGATGTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYOD1</td>
<td>AF027148</td>
<td>T</td>
<td>9898–9982bp</td>
<td>−35' −302</td>
<td>61</td>
<td>1.12</td>
<td>0.74</td>
<td>6FAM-CCTCCGACCGCGATTTGATC</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CALCA</td>
<td>X19349</td>
<td>T</td>
<td>1706–1806bp</td>
<td>−128' −28</td>
<td>60</td>
<td>0.72</td>
<td>0.53</td>
<td>6FAM-CGTGCTTGGTGGATG-TGTT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGMT</td>
<td>U59038</td>
<td>B</td>
<td>206–297bp</td>
<td>+392/+483</td>
<td>67</td>
<td>0.88</td>
<td>0.55</td>
<td>6FAM-CGTGCTTGGTGGATG-TGTT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIMP3</td>
<td>U33110</td>
<td>T</td>
<td>1051–1143bp</td>
<td>+1051/+1143</td>
<td>78</td>
<td>1.04</td>
<td>0.63</td>
<td>6FAM-ACCCGACCGCGATTTGATC</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>U02509</td>
<td>B</td>
<td>759–832bp</td>
<td>+7' +80</td>
<td>69</td>
<td>1.36</td>
<td>0.71</td>
<td>6FAM-ACCCGACCGCGATTTGATC</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARF</td>
<td>AF082338</td>
<td>T</td>
<td>5447–5515bp</td>
<td>−203' −135</td>
<td>78</td>
<td>2</td>
<td>0.89</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDH1</td>
<td>L34545</td>
<td>T</td>
<td>842–911bp</td>
<td>−180' −111</td>
<td>69</td>
<td>1.2</td>
<td>0.68</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDK2NB</td>
<td>D75756</td>
<td>T</td>
<td>350–430bp</td>
<td>+222/+302</td>
<td>64</td>
<td>1.44</td>
<td>0.82</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSTP1</td>
<td>M24485</td>
<td>T</td>
<td>1146–1245bp</td>
<td>−79' −21</td>
<td>75</td>
<td>2</td>
<td>0.93</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLL1</td>
<td>U26559</td>
<td>T</td>
<td>254–341bp</td>
<td>−662' −575</td>
<td>68</td>
<td>2.08</td>
<td>1.04</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTGS2</td>
<td>AF034206</td>
<td>T</td>
<td>67–924bp</td>
<td>−362' −217</td>
<td>62</td>
<td>0.8</td>
<td>0.71</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>THBS1</td>
<td>B28450</td>
<td>T</td>
<td>162–176bp</td>
<td>−636' −562</td>
<td>75</td>
<td>2.08</td>
<td>1</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTNNB1</td>
<td>X98448</td>
<td>T</td>
<td>583–668bp</td>
<td>−41V −332</td>
<td>75</td>
<td>1.92</td>
<td>1</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGACCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBI</td>
<td>L13190</td>
<td>T</td>
<td>1781–1900bp</td>
<td>−279' −160</td>
<td>68</td>
<td>2.24</td>
<td>1.27</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGATTCGCGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFB2B</td>
<td>U52240</td>
<td>T</td>
<td>256–333bp</td>
<td>−58' −19</td>
<td>75</td>
<td>2.08</td>
<td>0.81</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGATTCGCGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMYM</td>
<td>D00517</td>
<td>T</td>
<td>1052–1128bp</td>
<td>+301' +377</td>
<td>75</td>
<td>2.16</td>
<td>0.9</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGATTCGCGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIC-1</td>
<td>L41919</td>
<td>T</td>
<td>562–662bp</td>
<td>−55' +45</td>
<td>71</td>
<td>1.12</td>
<td>0.5</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGATTCGCGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTHFR</td>
<td>AF105977</td>
<td>T</td>
<td>123–219bp</td>
<td>+72' +45</td>
<td>57</td>
<td>0.56</td>
<td>0.54</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGATTCGCGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTB</td>
<td>Y00474</td>
<td>T</td>
<td>390–522bp</td>
<td>−1599' −1467</td>
<td>58</td>
<td>1.16</td>
<td>0.74</td>
<td>6FAM-CGTGCTTGGATTCGCGATT</td>
<td>GACGTTGCTGGATTCGCGAT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a There are two bases in our CDKN2A primers that differ from this GenBank sequence, because a preliminary high-throughput GenBank entry was the only available sequence at the time of our primer design. The correct primers should be the following: forward, TTGGAGTTGGTGGATGATT; and reverse, AGTGGATGGATGGATGATT. The modified base is in boldface type and underlined.

b The start site is not well defined.
ciliated DYS, tumor stage, and so forth, were used as the nominal variables. In the follow-up study of hypermethylation in IM and the presence of associated DYS and/or carcinoma, the IM samples were further dichotomized at 1 or fewer, versus two or more Class A genes methylated. A Fisher’s exact test was then used to determine statistical significance.

RESULTS

CpG Island Hypermethylation and the Progression of EAC. We analyzed the methylation status of a panel of CpG islands associated with 19 different genes and of one non-CpG island sequence for a total of 20 gene loci by the quantitative, high-throughput MethyLight assay (18, 19). The efficiencies of our methylation reactions were controlled for in each analysis by including unmethylated control DNA and methylated control DNA (14). The 20 genes were selected for their known involvement in carcinogenesis or because they have been shown to be methylated in other tumors (Table 1; Refs. 1, 6, 15, 23–39). We included a region located in the MTHFR gene as a control for a single-copy sequence that does not satisfy the criteria of a CpG island. CpG dinucleotides outside of an island are presumably normally methylated, unlike CpG dinucleotides within CpG islands.

Fig. 1 illustrates the quantitative methylation data of the 20 genes from our screen of 84 tissue specimens from 31 patients with different stages of Barrett’s esophagus and/or associated adenocarcinoma. There is a general increase in the frequency and in the quantitative level of CpG island hypermethylation at progressively advanced stages of disease. However, the propensity for aberrant methylation of the genes is not uniform. Genes differ both in their frequency and in their levels of hypermethylation in various tissues. Accordingly, genes can be grouped by their methylation behavior, as indicated in Fig. 1. This allows for a visual assessment of concordant methylation of the different genes. We provide a rationale for each of the gene classes in the following section.

Epigenetic Gene Classes. The analysis of combined behavior of genes with different levels of DNA methylation will lead to a bias of the group behavior toward genes with quantitatively high levels of DNA methylation. For instance, the mean values for group B for most of the tumor samples would be driven primarily by the TIMP3 values, because this gene tends to have higher levels of methylation than the other two genes in this group (see Fig. 1). To equalize the quantitative impact of methylated genes within each class, we collapsed the methylation values that were used to generate Fig. 1 into a binary variable with a dichotomization point of 4 PMR (see “Materials and Methods”). This dichotomization moderates the effect of highly methylated genes, simplifies cross-gene comparisons of methylation frequencies, as shown in Fig. 2, and allows the calculation of class averages of methylation frequencies, as shown in Fig. 3. We based the 4-PMR-dichotomization point on previous work (13), and on its ability to discriminate between the different tissue types, as shown in Fig. 1.

Of the panel of 20 genes, the most informative genes were those with an intermediate frequency of hypermethylation (ranging from 15% (CDKN2A) to 60% (MGMT) of the samples above the 4-PMR-methylation cutoff). This group could be further subdivided into three classes according to the absence (A) or presence (B) of methylation in normal esophageal mucosa and stomach, or the infrequent methylation of normal esophageal mucosa accompanied by methylation in all normal stomach samples (C). The other genes were less informative, because the incidence of hypermethylation was either very infrequent (Class D), completely absent (Class E), or ubiquitous (Classes F and G), regardless of tissue type (Figs. 1, 2, and 3).

Class A consists of the genes CDKN2A, ESR1, and MYOD1 (Figs. 1, 2, and 3). There is a statistically significant difference in the methylation frequency of ESR1 (P = 0.0001) and MYOD1 (P = 0.0038) of NE, as compared with IM tissue, but not for CDKN2A (P = 0.097). The frequency of CDKN2A methylation increases significantly in the more advanced stages of the adenocarcinoma (T) (P < 0.0001). Class B consists of the genes CALCA, MGMT, and TIMP3. In contrast to Class A, this class exhibits methylation in the NE mucosa and stomach (S) tissue (Figs. 1 and 2). Only TIMP3 shows a significant difference in methylation frequency between the NE and the IM (P = 0.0074).

Class C consists of the gene APC, which is, in contrast to Classes A and B, methylated in all of the normal stomach samples (Figs. 1 and 2). This confirms our previous documentation of APC methylation in normal stomach tissue (14). The mechanism that protects APC from methylation in the normal esophageal tissues but not in normal stomach tissues is not clear. Class D consists of the genes ARF, CDH1, CDKN2B, GSTP1, MLH1, PTGS2, and THBS1, which are infrequently methylated (Figs. 1 and 2). There is a slight increase in the frequency of this class of genes in adenocarcinoma, but this does not approach statistical significance (Fig. 3). Interestingly, with the exception of PTGS2, which has not yet been investigated in other systems, the remaining Class D genes are frequently hypermethylated in other tumor types (Table 2). Class E genes (CTNNB1, RB1, TGFB2, and TMS1) are unmethylated at each stage in the progression of EAC. Similar to most Class D genes, RB1 and TGFB2 have been found to be hypermethylated in other tumors types (Refs. 1, 15, 36, 37; Table 1).

It should be noted that all of the samples scored positive for DNA input as measured by the control gene (ACTB). Therefore, the lack of detectable DNA methylation cannot be attributed to a lack of input DNA. We verified that the control reaction was sufficient in each sample such that a level as low as 1 PMR for a given gene could be detected. The integrity and specificity of all of the methylation reactions was confirmed using in vitro methylated human DNA.

The Class F gene (HIC1) is completely methylated, regardless of tissue type (Figs. 1 and 2). HIC1 is commonly methylated in other types of cancers (1, 15), and has been shown to be methylated in normal breast ductal tissue and bone marrow samples of breast cancer and acute myelogenous leukemia patients, respectively (6, 40). Nevertheless, the finding of ubiquitous methylation of a CpG island in normal tissues was unexpected. Therefore, we confirmed the validity of the HIC1 MethyLight results using a different technique (HpaII-PCR; Ref. 41; data not shown). Interestingly, the ubiquitous HIC1 methylation pattern is similar to the non-CpG island MTHFR control (Class G); however the percentage of methylated molecules is quantitatively higher (Fig. 1).

Epigenetic Profiles of EAC Progression. Each tissue type shows a unique epigenetic profile that changes during disease progression (Fig. 3A). Classes A, B, and C are methylated at a significantly higher frequency in IM tissue than in normal esophageal mucosa (NE; Fig. 3, A and B). Furthermore, the transition from IM to DYS or malignancy (T) is associated with an additional increase in Class A methylation (Fig. 3, A and B). The lack of a significant difference between DYS and adenocarcinoma for any of the gene classes or when all of the 19 genes are combined (Fig. 3, A and B) suggests that most of these abnormal epigenetic alterations occur early in the progression of EAC.

Hypermethylation and EAC Tumor Grade and Stage. We investigated whether the grade or stage of an EAC correlates with a higher frequency of CpG island hypermethylation. Moderately differentiated tumors have significantly less frequent Class A methylation compared with poorly differentiated tumors (P = 0.045; data not shown). Fig. 4A shows that there is a significantly higher mean number of Class A genes methylated in stage II, III, and IV tumors.
relative to less advanced, stage I tumors. The differences between stage I tumors and stage II, III, and IV tumors did not reach statistical significance for any of the other classes.

**Methylation of Premalignant Tissues with or without Associated DYS.** The occurrence of CpG island hypermethylation in some cases of IM for Class A and some cases of normal esophageal mucosa for Class B raises the question as to whether these methylation events represent normal methylation patterns in these nondysplastic tissues, or whether they reflect methylation changes that predispose to further progression. In the latter case, one would expect to find a higher frequency of such CpG island hypermethylation in these tissues in patients who have already undergone further disease progression. Therefore, we have investigated whether this was the case in our study. Patients were divided based on whether or not they had Barrett’s esophagus (IM) as their most advanced stage of disease (Fig. 1, NO) or whether they had associated DYS and/or adenocarcinoma present in a different region of the esophagus (Fig. 1, YES). We found that the frequency of Class B methylation in the NE is indeed significantly higher in patients with associated DYS/tumor ($P = 0.0037$; Fig. 1). In addition, we found that Class A methylation is more frequent in IM samples from patients with concurrent DYS or

---

**Fig. 1.** Methylation analysis of a panel of 20 genes in tissues from 31 patients with Barrett’s esophagus and/or adenocarcinoma. Methylation analysis was performed using the MethyLight assay (18, 19). The percentage of fully methylated molecules at a specific locus (PMR) was calculated by dividing the GENE:ACTB ratio of a sample by the GENE:ACTB ratio of SssI-treated sperm DNA and multiplying by 100. □, samples with less than 4 PMR; □, 4–20 PMR; □, 21–50 PMR; □, samples with >50 PMR. On the left, the tissue types. On the right, numbers (1, 2, 3, and 4), TNM tumor staging; YES, presence of distally located DYS and/or adenocarcinoma in the patient; NO, absence of distally located DYS and/or adenocarcinoma in the patient. In the columns: N, an analysis for which the control gene ACTB did not reach sufficient levels to allow the detection of a minimal value of 1 PMR for that methylation reaction in that particular sample.

**Fig. 2.** Percentage of samples methylated for each gene by tissue type. The data were dichotomized with ≥4 PMR designated as methylated and <4 PMR as unmethylated. The genes are grouped according to their respective gene classes (A–G) as shown in Fig. 1. n, the number of samples analyzed for each tissue.
cancer, than in IM samples from patients without any evidence of further progression \( (P < 0.0001; \text{Fig. 1}) \). A potential criticism of this analysis is that the same set of samples was used to delineate the class of genes as was used to test the association with a clinical parameter. Therefore, we performed a follow-up study of 20 additional cases of IM, entirely independent of this first data set. Our initial study had revealed that all of the IM samples that were associated with further disease progression had at least two Class A genes methylated, whereas all of the IM samples without associated DYS or adenocarcinoma did not show any methylation of Class A genes (Fig. 1). Therefore, we defined two or more Class A genes methylated as an indicator of increased risk for the presence of associated DYS or cancer. The data from our first series gives a \( P = 0.0048 \) in a Fisher's exact test of this association (Fig. 5 \( A \)). Our new series of 20 independent cases gives a \( P = 0.018 \) (Fig. 5 \( B \)). We conclude that the positive association between hypermethylation of Class A genes and the presence of associated DYS or cancer is significant. It should be noted that the IM samples without associated DYS in this follow-up study show a low frequency of samples with at least two genes methylated, which is in contrast to the absence of methylation in the first study (Figs. 1 and 5). This may be attributed to the fact that the samples in the second series were microdissected from paraffin sections. Therefore, there is a lower background of unmethylated stromal cells in the sample. In this case, the methylation signal is not as diluted by other normal cells, and, consequently, the ratio of methylated molecules:total DNA may rise above the 4 PMR threshold. Alternatively, dysplastic or malignant tissue may have been missed during the endoscopic survey in some of the cases scored as free of further disease progression because of the sampling limitations of endoscopy. This is a well-documented problem in the detection of EAC (42).

CIMP Analysis. We investigated whether EAC tumors exhibit a CIMP. CIMP tumors are a distinct group of tumors that are defined by a high degree of concordant CpG island hypermethylation of genes exclusively methylated in cancer, or type C genes (16, 17). In our analysis, Class A genes most exemplify the type C genes because they lack methylation in the normal tissues. We examined the distribution of the number of Class A genes methylated and found that the frequency of genes methylated in the adenocarcinoma tissue does not show the expected bimodal distribution of CIMP (Fig. 6; Ref. 17). Similar results are observed when we also include Class D genes (Fig. 6), which also exhibit type C methylation, and when we combine Classes A, B, C, and D genes (Fig. 6). Classes E and F genes are not included because they do not exhibit any methylation variation between the different tissue
types. We did find a single sample with 10 of 14 Class A-D genes methylated (Fig. 1, Case 3, and Fig. 6). However, this sample only stands out when we include Class B genes, which are methylated in normal esophageal mucosa and, therefore, do not satisfy the definition of type C genes that constitute the CIMP phenotype (16, 17). Therefore, we tentatively conclude that there is no clear evidence of a separate group of CIMP tumors in our study of EAC, as previously defined for colorectal and gastric cancer (16, 17). We do find, however, that CpG island hypermethylation in EAC does occur across multiple loci in a given sample. Furthermore, the number of loci hypermethylated in a single sample increases as the disease progresses through different histological stages (Fig. 6). The bimodal distributions seen in IM tissues (Fig. 6) can be fully attributed to the concurrent association with DYS or cancer described above.

DISCUSSION

We have used a high-throughput, fluorescence-based methylation assay (MethyLight) to examine the hypermethylation patterns of 19 CpG islands and 1 non-CpG island during the progression of EAC. We analyzed this relatively large number of genes with the goal of identifying diagnostic epigenetic changes during disease progression and to investigate whether these epigenetic changes show evidence of group behavior. Our purpose was not to draw conclusions about the functional significance of individual methylation events to the disease process or to infer effects on gene expression levels. The presence of substantial amounts of normal tissue in our frozen specimens and the use of formalin fixation in our microdissected samples hinder the assessment of the gene inactivation effects of the CpG island hypermethylation that we have detected (14). The analysis of abnormal DNA hypermethylation offers an advantage over gene expression analysis in that it has greater sensitivity in the presence of contaminating normal cells. To our knowledge, this is the most comprehensive methylation survey ever performed on a system with so many distinct histological stages of disease progression.

We dichotomized the methylation data to equalize the quantitative impact of methylated genes within each class, simplifying cross-gene comparisons of methylation frequencies. However, the dichotomization point does not significantly affect the statistics or alter the conclusions. For instance, there is still a statistically significant difference in the mean percentage of genes methylated (of 19 genes) between the normal esophageal mucosa and the IM, DYS, and adenocarcinoma tissues when the data are dichotomized at 10 PMR. In addition, all of the statistically significant findings of the NE and IM methylation frequency with or without associated DYS remain significant at a dichotomization point of 10 PMR, instead of 4 PMR. It is important to note that 4 PMR is not comparable with a 4% methylation level of a single CpG dinucleotide. Rather, it indicates that in this sample, 4% of the DNA molecules had complete methylation at all of the CpG dinucleotides covered by the three MethyLight primers (usually about eight CpGs). The nature of the MethyLight assay is such that it is oblivious to any other methylation patterns that may be present (19). Therefore, 4 PMR is likely to represent a higher mean level of methylation than 4%. The extensively methylated molecules that are assayed by MethyLight are likely to represent alleles that have been completely silenced by CpG island hypermethylation, although this was not investigated in this study.
HYPERMETHYLATION IN EAC

This genomic approach to DNA methylation changes (as opposed to a functional approach) leads us to several interesting findings and conclusions. It is now clear that DNA hypermethylation is an early epigenetic alteration in the multistep progression of EAC. The pre-malignant IM is already significantly more methylated than the normal tissue. We report, for the first time, frequent hypermethylation of five additional genes in this tumor system: MYOD1, MGMT, CALCA, TIMP3, and HIC1. The methylation observed for MGMT, TIMP3, and HIC1 in normal tissues may be attributed to the region of the gene in which we analyzed methylation levels (43–45). These three genes were analyzed at CpG islands located at, or downstream of, the transcription start site (Table 2). However, this does not account for the CALCA methylation that we observed because we analyzed the promoter region of this gene. Low levels of CALCA methylation has been previously reported in normal bone marrow samples of acute myelogenous leukemia patients (6), which suggests that this locus may have a higher propensity to be methylated in the normal tissues of cancer patients.

It is of particular interest to note that dysplastic tissues are more frequently methylated than stage I tumors for both Class A (P < 0.0001) and B (P = 0.0174; Fig. 1). This is similar to the finding of genetic abnormalities (loss of heterozygosity, deletions, and mutations) present in Barrett’s esophagus with high-grade DYS but not present in the adjacent invasive EAC (10). Because stage II-IV tumors appear to be methylated in Class A genes at a similar frequency as DYS, this suggests that stage I tumors may actually evolve from an origin different from that of the dysplastic tissue and higher-staged tumors or may diverge after DYS independently from stage II-IV tumors during clonal expansion. Alternatively, stage I tumors could undergo a transient reversal of hypermethylation, although we consider that to be rather unlikely. Tumor development in Barrett’s esophagus is proposed to evolve clonally through the linear multistep pathway of metaplasia-DYS-tumor (11). However, the occurrence of genetic and, now, epigenetic alterations in a nonlinear order suggests that the clonal evolution of EAC is more complex than originally predicted (10). A similar observation has been described for different stages of bladder tumors (46).

We did not find clear evidence, aside from one tumor with 10 genes methylated, for a separate cluster of tumors with extensive concordant methylation, which would be indicative of a CIMP. Similar results are obtained even if we examine only type C genes, as defined for CIMP (methylated in cancer, not methylated in normal tissues; Refs. 16, 17). Interestingly, the type C genes in EAC differ from those described for colorectal cancer (16, 17). For example, ESR1 is classified as a type A gene (defined as methylated in aging normal tissues) rather than a type C gene in colorectal cancer, because it is frequently methylated in the normal colonic epithelium of aging individuals (16, 17). However, in EAC, ESR1 clearly behaves like a type C gene. This may be attributed to the difference in the technology used to measure hypermethylation or more likely may be attributable to differences in tissue types.

There is clearly a tissue-specific and tumor-specific propensity for particular genes to become hypermethylated. For instance, APC is hypermethylated in normal stomach, but not in normal esophageal mucosa. The tumor-specificity of hypermethylation is illustrated by the lack of detectable methylation of the two Class E genes TGFBR2 and RB1, which are frequently hypermethylated in gastric and lung tumors, and retinoblastoma tumors, respectively (3, 36, 37). The tumor-specificity of CpG island hypermethylation suggests that there may be tissue-specific trans-acting factors that modulate methylation changes of these CpG islands during tumorigenesis and that differ between EACs and other tumor types. Alternatively, there may be a lack of selective advantage to the silencing of these genes in EACs by DNA methylation. There are two scenarios in which this would be the case. One is if the gene in question has been inactivated by a different, genetic mechanism, rendering hypermethylation of no further selective advantage. The other is if the gene does not play a role in tumor suppression in this particular tumor system.

Although alterations in DNA methylation changes are common events in tumorigenesis, the underlying mechanism is unclear. We have shown previously in colorectal tumors that abnormal methylation is not attributable to a mere up-regulation of the DNA methyl-transferase genes, which suggests that other major players are involved (18). Our data do provide some first glimpses into the process underlying these abnormal methylation changes. We find that different, functionally unrelated, genes can behave in distinct classes with respect to their methylation changes within various tissues of EAC progression. The CpG island hypermethylation does not appear to be a random, stochastic process (although there is a stochastic component), but rather a step-wise process that involves multiple, distinct groups of alterations. This is consistent with the existence of several different mechanisms that protect against CpG island hypermethylation. In this scenario, the concerted changes seen at different CpG islands would be the result of the loss of a different type of protective element at different stages of disease progression. This finding does not appear to be dependent on the location of the CpG island relative to the gene, because both promoter and internal CpG islands were observed in all of the gene classes. We also examined the structural features of these CpG islands by analyzing the percentage of GC content, the observed/expected CpG ratio, and the CpG-GpC ratio, and we found no association with gene class (Table 2).

We show that the IM or NE samples themselves, with or without associated DYS or cancer, were histologically indistinguishable, yet molecularly distinct. NE and IM samples derived from individuals with concurrent, distally located DYS or malignancy show a statistically higher incidence of CpG island hypermethylation. We confirmed this finding in the IM tissues in a completely independent study (Fig. 5B). This suggests that epigenetic markers, particularly Class A and B genes, may be used as disease screening tools and perhaps ultimately as predictive markers for the progression of more advanced-stage disease. However, because this was not a longitudinal study, we cannot distinguish whether the methylation seen in the preneoplastic tissues represents a predisposing event in the field of cells that give rise to the DYS, or whether the methylation in the preneoplastic tissue develops concurrently with the DYS or adenocarcinoma. Regardless, these methylation profiles could be of profound significance in the early detection of this disease. A molecular diagnostic approach using normal and/or premalignant tissues that might identify patients with cancer or at elevated risk for developing cancer would provide an opportunity for early intervention. Furthermore, a benefit of using CpG island hypermethylation as a diagnostic marker is that it can easily be detected in a field of normal cell contamination as a gain of signal, unlike loss of gene expression, loss of heterozygosity, and deletion analysis, which are difficult to resolve in a sample with contaminating normal cells. A prospective longitudinal study should help reveal whether these epigenetic alterations in normal or early-stage tissues are predictive of imminent dysplastic/malignant disease. A concern in this part of our study is the potential for the presence of contaminating tumor cells in the NE and IM tissue. If the methylation that we observe in nondysplastic tissues were indeed attributable to admixture of dysplastic or malignant tissues, then the level of contamination would have to be at least 4 PMR, because we use a 4-PMR dichotomization point. Although the presence of contaminating dysplastic/tumor cells can never be fully excluded, this is unlikely to be the case, because not only were the NE and IM samples separated by large distances (≥10 cm in the case of
the normal tissue) or microdissected (IM samples in the second study) from the malignant and dysplastic areas, but each sample was also very carefully reviewed by a pathologist for any signs of dysplasia or DYS in both of the studies.

In summary, the 19 CpG islands segregate into six classes of epigenetic patterns in the various tissue types. Each class undergoes unique epigenetic changes at different steps of disease progression of EAC. We propose that these unique epigenomic profiles arise from an interplay between gene-specific cis-acting factors and tissue- and tumor-specific trans-acting factors. Additional studies such as this one will help to clarify this picture and identify clusters of CpG islands that behave concordantly in different systems. The next step will then be to identify the molecular mechanisms and factors affecting the various CpG island clusters.

REFERENCES


