

Epigenomic program of Barrett's-associated neoplastic progression reveals possible involvement of insulin signaling pathways

Dear Editor

Although the molecular basis of esophageal adenocarcinoma (EAC) and its precursor, Barrett's esophagus (BE), have been studied, further insights are needed into early premalignant events. In the current study, we comprehensively identified epigenetically altered genes at various time points during neoplastic progression in BE. We compared global DNA methylation profiles of two groups of BE patients: 1) patients who later progressed to high-grade dysplasia (HGD) or EAC, termed 'progressors (P)'; and 2) those who did not progress to HGD or EAC, termed 'nonprogressors (NPs)'.

First, we conducted a comprehensive, genome-wide CpG island methylation analysis on a cohort of nine BE specimens. These specimens were obtained from two sets of Barrett's metaplasia patients with known outcomes: five were from Ps and four were from NPs. All the tissue specimens used in this study were obtained from patients who provided informed written consent under an approved protocol by the Institutional Review Board at the University of Maryland, School of Medicine, the Baltimore Veterans Affairs Medical Center and the Johns Hopkins University School of Medicine. Ps were defined as patients with index (i.e. pre-progression experimental study) biopsies showing either no dysplasia, indefinite for dysplasia, or low-grade dysplasia (LGD) at endoscopy performed at least 6 months before the diagnosis of either HGD or EAC. NPs were defined as patients who had undergone at least three surveillance endoscopic examinations with index biopsies but did not progress beyond LGD.

CpG island methylation analysis was conducted by combining two techniques: 1) methylated CpG island amplification (MCA) and 2) CpG island oligonucleotide microarrays. The MCA method was first used to enrich and amplify methylated DNA fragments (Toyota *et al.* 1999). These enriched methylated DNA fragments were then labeled and hybridized to 244K Human CpG Island microarrays (Agilent Technologies, Santa Clara, CA, USA). These

microarrays contain 244 000 probes corresponding to 27 800 CpG islands, covering ~21 megabases (MB) of the genome. Each array was co-hybridized to fully methylated DNA, labeled with Cy3, as a universal control, and to a single P or NP specimen-derived DNA, labeled with Cy5. Probe-based signal intensity data were normalized using the LOWESS-based method. Data obtained were expressed as ratios of Cy5 to Cy3, i.e. sample DNA to fully methylated control DNA. After removing results for fragments > 2500 bp in length, data on 34 396 SmaI/XmaI (S/X) target fragments remained. Each S/X fragment corresponded to a CpG island within either a gene promoter region, an intragenic region, or an extragenic region. Many promoters or genes were represented by multiple S/X fragments. We then took the mean signal intensity value of the four NPs and of the five Ps for each S/X fragment, and calculated the P/NP ratio for each S/X fragment. The most significant 25 of these fragments by *t*-test, sorted from the lowest to highest *P* value, are displayed in Table 1.

As can be seen in Table 1, among the top-ranked 25 loci, only three (*Pro_MMD2*, *Pro_ZNF358* and *Intra_F10*; in bold) were more methylated in Ps than in NPs (i.e. P/NP ratio > 1.0; second-to-last column). The remaining 22 loci were relatively *hypomethylated* in Ps vs NPs.

These results, for the first time, established a paradigm shift away from the traditional assumption that the predominant epigenetic change in BE progression is methylation-induced silencing of TSGs. Based on these findings, just the opposite appears to be the case: *hypomethylation* prevails, thus *activation* (presumably, at least in part) of *growth-promoting* genes seems to constitute the predominant epigenetic program. Several genes that we found to be hypomethylated in progressors do, indeed, bear out this theory: *IGF1R*, encoding the insulin-like growth factor 1 receptor, corresponded to the fourth most significantly different among 34 396 S/X fragments studied (*t*-test, *P* value = 0.00078) and was hypomethylated in Ps vs NPs (mean P/N = 0.585). This tyrosine kinase receptor

Table 1 The top 25 loci from methylation array data of BE progressors vs NPs (sorted by the lowest to highest P value). Target location_ annotation, location of the array probe (Intra, intragenic; Pro, within promoter; Extra, extragenic) and annotation of the loci; NP, nonprogressor; Mean, average of all NPs or all Ps; P, progressor; P/NP ratio, ratio of means of Ps to means of NPs; *P value: Student's t-test value of NP group vs P group. When this table was expanded further, the predominance of hypomethylated loci persisted (not shown due to space limitations)

Target location_ annotation	Nonprogressor					Progressor					P value		
	NP1	NP2	NP3	NP4	Mean	P1	P2	P3	P4	P5		Mean	P/NP ratio
<i>Intra_CCDC57</i>	9.45	7.21	8.57	8.17	8.35	1.47	1.33	1.01	2.83	3.10	1.95	0.23	0.00003
<i>Pro_MGC35308</i>	2.05	2.05	2.39	2.06	2.14	1.00	1.03	1.14	1.41	0.98	1.11	0.52	0.00007
<i>Pro_TLX3</i>	6.68	6.28	7.21	5.50	6.42	0.98	1.19	2.28	2.85	2.73	2.01	0.31	0.00007
<i>Pro_IGF1R</i>	0.91	0.85	0.85	0.97	0.90	0.60	0.58	0.46	0.42	0.55	0.52	0.58	0.00008
<i>Pro_ITGB8</i>	2.06	1.87	1.72	1.87	1.88	0.99	0.84	0.95	1.08	1.30	1.03	0.55	0.00008
<i>Pro_HOXB7</i>	1.62	1.51	1.60	1.71	1.61	0.93	0.99	1.14	1.21	1.14	1.08	0.67	0.00010
Extragenic	19.97	18.90	17.03	14.22	17.53	1.13	1.56	1.31	7.62	4.99	3.32	0.19	0.00011
<i>Pro_ZBTB7B</i>	1.39	1.39	1.49	1.65	1.48	0.71	0.75	0.78	1.00	0.88	0.83	0.56	0.00011
<i>Pro_MMD2</i>	0.84	0.77	0.74	0.85	0.80	1.04	1.09	1.14	1.04	1.04	1.07	1.34	0.00014
<i>Pro_JUND</i>	2.68	2.31	2.64	2.30	2.48	1.15	1.15	1.20	1.72	1.51	1.34	0.54	0.00015
<i>Pro_UBP1</i>	1.30	1.35	1.18	1.28	1.28	0.81	0.84	0.98	0.90	0.98	0.90	0.71	0.00016
<i>Pro_LOCS5565</i>	1.15	1.10	1.20	1.09	1.14	0.85	0.86	0.81	0.97	0.81	0.86	0.76	0.00017
<i>Pro_LAMA5</i>	1.26	1.23	1.39	1.07	1.24	0.47	0.45	0.61	0.71	0.54	0.56	0.45	0.00017
<i>Pro_ZNF358</i>	0.56	0.61	0.55	0.57	0.57	0.77	0.73	0.71	0.69	0.70	0.72	1.25	0.00018
<i>Pro_GPR177</i>	0.29	0.26	0.29	0.28	0.28	0.19	0.18	0.18	0.22	0.21	0.20	0.70	0.00018
Extragenic	3.51	2.75	3.43	3.27	3.24	1.22	1.15	1.27	1.56	1.71	1.38	0.43	0.00020
<i>Intra_BCL11B</i>	53.82	45.41	52.89	62.47	53.65	1.29	1.70	5.13	28.54	7.78	8.89	0.17	0.00020
<i>Intra_F10</i>	0.73	0.87	0.58	0.69	0.72	1.39	1.21	1.28	1.56	1.24	1.33	1.86	0.00021
<i>Pro_CKB</i>	1.37	1.32	1.36	1.31	1.34	1.05	1.09	1.17	1.18	1.09	1.12	0.83	0.00021
<i>Pro_TAF10</i>	1.96	2.11	2.27	2.06	2.10	0.70	0.63	0.79	1.28	1.16	0.91	0.43	0.00022
<i>Pro_ELAVL3-ZNF653</i>	2.04	1.92	1.72	1.74	1.85	0.95	1.05	1.16	1.35	1.06	1.11	0.60	0.00024
<i>Pro_WNK4</i>	0.81	0.82	0.84	0.80	0.82	0.59	0.70	0.60	0.59	0.61	0.62	0.76	0.00025
Extragenic	1.93	1.99	2.43	2.31	2.16	0.71	0.69	1.08	1.30	1.12	0.98	0.45	0.00026
<i>Pro_WWC1</i>	1.01	1.06	1.18	1.11	1.09	0.70	0.65	0.87	0.73	0.74	0.74	0.68	0.00026
<i>Pro_JARID 1B</i>	5.86	5.28	4.35	4.66	5.04	1.03	1.17	1.59	2.99	1.93	1.74	0.35	0.00027

is highly overexpressed in most malignant tissues, where it functions as an anti-apoptotic agent by enhancing cell survival (Ewing & Goff 2010). Similarly, *TLX3*, a *HOX11* gene family homeobox member encoding a DNA-binding nuclear transcription factor, is overexpressed in T-cell leukemias and associated with a poor prognosis (Attarbaschi *et al.* 2009). It was the third most significant ($P=0.000073$) and 0.313-fold hypomethylated in Ps vs NPs. *JUND*, another hypomethylated and the tenth most significant gene, is a well-known proto-oncogene and functional component of the AP1 transcription factor complex that has been proposed to protect cells from p53-dependent senescence and apoptosis (Weitzman *et al.* 2000). The methylation status of *TLX3*, *IGFIR* and *JUND* in the P and NP samples that were studied by microarray analysis was confirmed by real-time methylation-specific PCR (qMSP) assays, using loci-specific primers and probes (data not shown).

Next, using prediction analysis of microarrays (PAM), we constructed a classifier to distinguish Ps from NPs based on the complete set of methylation array data. Interestingly, PAM classified all the samples correctly except for P4, which was consistently misclassified as a NP. We then inspected the clinical data on patient P4. We noticed that the time interval between sample collection and progression in patient P4 was 40 months. Thus, P4 had been obtained before the pre-progression ‘window’ of 2 years, the earliest time point at which previously studied methylation marker levels have been shown to begin to change (Schulmann *et al.* 2005). In contrast, the remaining four progressors had pre-progression intervals ranging from 16 to 31 months, closer to or within the previously reported 24-month window.

We next inspected the complete methylation array data for methylation levels of specific genes in individual P and NP specimens. Results of this inspection are illustrated in Table 2. This table illustrates that many of the genes in P4 exhibited methylation levels intermediate between those of the other Ps and those of the NPs. For example, genes highlighted in orange were less hypomethylated in P4 than in the other progressors. Conversely, genes highlighted in blue were less hypermethylated in P4 than in the other progressors. By contrast, there were other genes for which P4 perfectly matched the remaining progressors. For example, genes highlighted in green were just as hypomethylated in P4 as in the other progressors; and genes highlighted in pink were just as hypermethylated in P4 as in the other progressors. Thus, some genes (the green and pink genes) appeared to be good early predictors of progression, while others (the orange and blue genes)

Table 2 Representative SAM-processed data, sorted by corrected P (i.e. SAM q)-value, on genes, progressors, nonprogressors, and ‘outlier’ index sample P4. This table illustrates that in P4, the sample with the longest pre-progression interval, many genes exhibited methylation levels intermediate between the other progressors and the nonprogressors. Some genes are either less hypomethylated or less hypermethylated in P4 than in the other progressors, while others are similarly hypo- or hypermethylated in all progressors

Target location_ annotation	Nonprogressor					Progressor					SAM score	q value
	NP1	NP2	NP3	NP4	P1	P2	P3	P4	P5			
<i>Pro_DSCR1L2</i>	2.36	2.22	2.49	2.06	1.13	1.21	1.40	1.30	1.26	10.63	0	
<i>Pro_CPEB1</i>	5.86	4.38	6.23	6.02	0.79	0.70	1.58	1.95	1.43	−9.42	0	
<i>Pro_LY6H</i>	2.28	3.18	3.23	2.41	0.99	1.00	0.73	0.76	0.70	−8.35	0	
Extragenic	19.97	18.90	17.03	14.22	1.13	1.56	1.31	7.62	4.99	−7.74	0	
<i>Pro_IGF1R</i>	0.91	0.85	0.85	0.97	0.60	0.58	0.46	0.42	0.55	−7.69	0	
<i>Pro_CYP2E1</i>	0.91	0.94	0.79	0.90	1.14	1.19	1.24	1.16	1.08	6.35	1.38	
<i>Intra_SCRIPT2</i>	0.76	0.83	0.51	0.66	1.31	1.42	1.26	1.06	1.31	6.38	1.38	
<i>Pro_LOC388335</i>	0.46	0.58	0.42	0.57	1.17	1.33	1.19	0.88	1.36	6.49	1.38	
<i>Pro_RNF126</i>	0.83	0.81	0.74	0.76	0.94	0.97	1.02	0.97	0.94	6.97	1.38	

tended to change their methylation levels *later*, i.e. temporally closer to the progression endpoint.

We reasoned that ‘late-demethylator’ genes or loci (such as *Pro_RAX*) could represent hypomethylation events occurring closer to the progression endpoint, while ‘early-demethylator’ genes or loci (such as *Pro_ITGB8*) could become hypomethylated at longer time intervals prior to progression. Similarly, ‘late-methylator’ genes or loci (such as *Intra_PSCD3*) were seen to potentially represent hypermethylation events occurring later in Barrett’s neoplasia (BN), i.e. closer to the progression endpoint, while ‘early-methylator’ genes or loci (such as *Pro_WDR5*) were hypothesized to become hypermethylated at longer time intervals before progression. A graphic illustration of this concept is provided in Fig. 1.

By using comprehensive methylation microarrays, we were able to seek and find genes that became hypomethylated earlier, even as early as 40 months, prior to progression. Similarly, although P4 was an ‘outlier’ for loci that became hypermethylated during progression (Fig. 1b), methylation microarrays enabled us to find genes that became hypermethylated in P4 at 40 months prior to progression (*pink* gene, Fig. 1b).

Next, we attempted to verify this late- vs early-methylator/demethylator hypothesis by applying hierarchical clustering to methylation array data from

our five progressor specimens, using Statistica (v. 7.0; Statsoft, Tulsa, OK, USA). Hierarchical clustering revealed two principal branches in these data: one main branch containing samples P1, P2, P3, and P5; and the other main branch consisting entirely of the outlier sample P4 (data not shown). Thus, P4 *per se* appeared, on a global basis, to constitute a category that differed markedly from the remaining four progressors.

Furthermore, we sought to identify ontologic gene groupings. We applied GOstat (<http://gostat.wehi.edu.au/cgi-bin/goStat.pl>) to determine which gene ontology terms were over- or under-represented in genes whose promoter regions were differentially methylated in the outlier specimen P4 vs the remaining four Ps. Differential methylation in this analysis was defined by the following two criteria: 1) the difference in the \log_2 intensity ratio of P4 vs the mean of P1, P2, P3, and P5 had to be above 0.5 (i.e. equivalent to a 1.4-fold non-log change); and 2) the \log_2 intensity ratio of the outlier P4 had to lie outside the 99% confidence intervals of the mean for samples P1, P2, P3, and P5. GOstat then identified Gene Ontologies that were significantly over-represented in P4 (the ‘orange’ genes) relative to all genes whose promoter regions were analyzed in our methylation array analysis. Several of these ontologic categories related to organ development, system development, or multicellular organismal development.

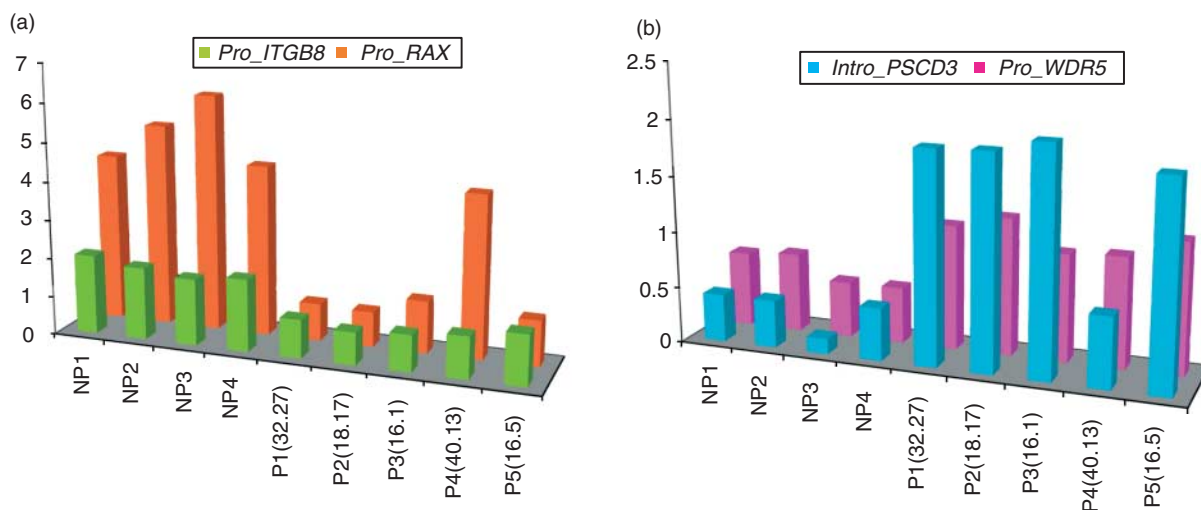


Figure 1 (a) Sample (‘slice’) of the temporal epigenomic program of Barrett’s pre-progression, focusing on hypomethylation. X-axis, pre-progression interval of progressor samples or nonprogressors (NPs). Y-axis, methylation index. Orange bars (rear) represent genes, such as *RAX*, which are consistently and significantly less methylated in progressors than in NPs, *except* for the outlier sample P4. Green bars (front) signify genes, such as *ITGB8*, that are hypomethylated in *all* progressors relative to NPs, *including* sample P4 (interval to progression = 40.13 months). (b) Sample (‘slice’) of the temporal epigenomic program of Barrett’s pre-progression, focusing on hypermethylation. X-axis, pre-progression interval of progressor samples or NPs. Y-axis, methylation index. Blue bars (front) identify loci, such as *Intra_PSCD3*, which are consistently and significantly more heavily methylated in progressors than in NPs, *except* for the outlier sample P4. Pink bars (rear) denote genes, such as *WDR5*, that are hypermethylated in *all* progressors relative to NPs, *including* sample P4 (interval to progression = 40.13 months).

As a caveat, we did not employ multiple testing correction, thus some of the observed differences could have occurred by chance. Moreover, distinguishing characteristics of sample P4 besides its pre-progression interval, such as smoking or therapeutic exposure, could have contributed to the uniqueness of this sample. Nevertheless, based on the above findings, it is tempting to speculate that during the transition from early to late pre-progression, the epigenomic program ‘switches’ to become dominated by aberrant organ development. Another insight into this process, as mentioned above, is the observed dominance of hypomethylation over hypermethylation in the epigenomic program of BE pre-progression.

Rachana Agarwal^{1*}

Zhe Jin^{3*}

Jian Yang¹

Yuriko Mori¹

Jee Hoon Song¹

Sahil Kumar⁴

Masato Sato⁵

Yulan Cheng¹

Alexandru V Olaru¹

John M Abraham¹

Amit Verma⁶

Stephen J Meltzer^{1,2}

¹Division of Gastroenterology, Department of Medicine and ²Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

³Department of Pathology, The Shenzhen University School of Medicine, Shenzhen, Guangdong, People’s Republic of China

⁴All India Institute of Medical Sciences, New Delhi, India

⁵Tohoku University School of Medicine, Miyagi, Japan

⁶Department of Medicine, Albert Einstein College of Medicine, Bronx, New York, USA

(Correspondence should be addressed to R Agarwal, Division of Gastroenterology, Department of Medicine, Johns Hopkins University School of Medicine, 1503 East Jefferson Street, Room 101, Baltimore, Maryland 21287, USA;

Email: ragarwa@jhmi.edu.

S J Meltzer,

Division of Gastroenterology, Department of Medicine, Johns Hopkins University School of Medicine, 1503 East Jefferson Street Room 112, Baltimore, Maryland 21287, USA;

Email: smeltzer@jhmi.edu)

*(R Agarwal and Z Jin contributed equally to this work)

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by the National Institute of Diabetes and Digestive and Kidney Disease (Grant no.: DK087454).

References

- Attarbaschi A, Pisecker M, Inthal A, Mann G, Janousek D, Dworzak M, Potschger U, Ullmann R, Schrappe M, Gadner H *et al.* 2009 Prognostic relevance of TLX3 (HOX11L2) expression in childhood T-cell acute lymphoblastic leukaemia treated with Berlin–Frankfurt–Munster (BFM) protocols containing early and late re-intensification elements. *British Journal of Haematology* **148** 293–300. (doi:10.1111/j.1365-2141.2009.07944.x)
- Ewing GP & Goff LW 2010 The insulin-like growth factor signaling pathway as a target for treatment of colorectal carcinoma. *Clinical Colorectal Cancer* **9** 219–223. (doi:10.3816/CCC.2010.n.032)
- Schulmann K, Sterian A, Berki A, Yin J, Sato F, Xu Y, Olaru A, Wang S, Mori Y, Deacu E *et al.* 2005 Inactivation of p16, RUNX3, and HPP1 occurs early in Barrett’s-associated neoplastic progression and predicts progression risk. *Oncogene* **24** 4138–4148. (doi:10.1038/sj.onc.1208598)
- Toyota M, Ho C, Ahuja N, Jair KW, Li Q, Ohe-Toyota M, Baylin SB & Issa JP 1999 Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Research* **59** 2307–2312.
- Weitzman JB, Fiette L, Matsuo K & Yaniv M 2000 JunD protects cells from p53-dependent senescence and apoptosis. *Molecular Cell* **6** 1109–1119. (doi:10.1016/S1097-2765(00)00109-X)