Peripheral Blood Immune Cell Methylation Profiles Are Associated with Nonhematopoietic Cancers

Devin C. Koestler1, Carmen J. Marsit1,2, Brock C. Christensen1,2, William Accomando4, Scott M. Langevin3,4, E. Andres Houseman5, Heather H. Nelson6, Margaret R. Karagas1, John K. Wiencke7, and Karl T. Kelsey3,4

Abstract

Background: Blood leukocytes from patients with solid tumors exhibit complex and distinct cancer-associated patterns of DNA methylation. However, the biologic mechanisms underlying these patterns remain poorly understood. Because epigenetic biomarkers offer significant clinical potential for cancer detection, we sought to address a mechanistic gap in recently published works, hypothesizing that blood-based epigenetic variation may be due to shifts in leukocyte populations.

Methods: We identified differentially methylated regions (DMR) among leukocyte subtypes using epigenome-wide DNA methylation profiling of purified peripheral blood leukocyte subtypes from healthy donors. These leukocyte-tagging DMRs were then evaluated using epigenome-wide blood methylation data from three independent case-control studies of different cancers.

Results: A substantial proportion of the top 50 leukocyte DMRs were significantly differentially methylated among head and neck squamous cell carcinoma (HNSCC) cases and ovarian cancer cases compared with cancer-free controls (48 and 47 of 50, respectively). Methylation classes derived from leukocyte DMRs were significantly associated with cancer case status (P < 0.001, P < 0.03, and P < 0.001) for all three cancer types: HNSCC, bladder cancer, and ovarian cancer, respectively and predicted cancer status with a high degree of accuracy (area under the curve [AUC] = 0.82, 0.83, and 0.67).

Conclusions: These results suggest that shifts in leukocyte subpopulations may account for a considerable proportion of variability in peripheral blood DNA methylation patterns of solid tumors.

Impact: This illustrates the potential use of DNA methylation profiles for identifying shifts in leukocyte populations representative of disease, and that such profiles may represent powerful new diagnostic tools, applicable to a range of solid tumors. Cancer Epidemiol Biomarkers Prev; 21(8); 1293–302. ©2012 AACR.

Introduction

Over the past decade, major advances have been made toward the understanding of pathogenesis by examining DNA methylation signatures between cancer and cancer-free subjects. This has revealed profoundly aberrant patterns of DNA methylation in cancer and has contributed to a growing understanding disrupted cellular functioning through epigenetic mechanisms (1–4). Much of the research in cancer epigenetics, however, has focused on examining profiles of methylation within the target cells of the tumor tissue itself (5–9) and only recently, has attention been directed toward examining methylation signatures in blood for nonhematopoietic malignancies (10–15). In the first large-scale epigenome-wide study of peripheral blood DNA methylation, profiles of blood-derived DNA methylation were shown to predict active ovarian cancer with considerably high sensitivity and specificity, area under the curve (AUC) = 0.80 (11). Subsequent studies involving epigenome-wide assessment of peripheral blood methylation have revealed similarly impressive prediction performance; AUC = 0.70 in a study of bladder cancer (13), AUC = 0.73 in head and neck squamous cell carcinoma (HNSCC; ref. 15), and in a 2-phase study of pancreatic cancer, AUC values of 0.85 and 0.76 in phases 1 and 2, respectively, for differentiating cases and controls (14). These findings suggest that assessment of DNA methylation in peripheral blood of patients with cancer could offer important new insights into the pathophysiology of cancer while also serving as a promising new avenue for noninvasive cancer detection and
diagnostics. Despite these highly significant findings, however, the biologic mechanisms underlying these clinically important results remains unclear.

Research examining peripheral blood DNA methylation have, through post-hoc bioinformatic pathways analyses, suggested that profiles of DNA methylation alteration associated with cancer are overrepresented with genes involved in immune system modulation (11–15), and so alterations in blood-derived DNA methylation may reflect changes to the white blood cell (WBC) composition in peripheral blood as a mediator or consequence of tumorigenesis (11). To date, no studies have conclusively, experimentally, evaluated whether or not the observed differences in DNA methylation profile represent differences in the underlying population of cells examined. Such a mechanistic understanding of the observed associations is necessary for applying these novel molecular diagnostic strategies in clinical practice.

Because of the potential clinical use of new epigenetic biomarkers for early detection of cancers (16, 17), we sought to address this mechanistic gap. We hypothesized that epigenetic signatures in blood that differentiate cancer cases from controls arise as a result of specific immune responses represented by shifts in leukocyte populations. To address this hypothesis, we first examined epigenome-wide DNA methylation in magnetic antibody sorted, normal human peripheral blood leukocyte subtypes to discern differentially methylated regions (DMR) that differentiate leukocyte subtypes. These leukocyte-tagging DMRs were then investigated using epigenome-wide blood methylation data from 3 independent case-control studies of different cancers: a HNSCC data set (15), an ovarian cancer data set (11), and a bladder cancer data set (13). Through these analyses, we provide a more thorough mechanistic understanding for the observed associations between peripheral blood DNA methylation and the presence of solid tumors.

Materials and Methods
Study population
Sorted, normal, human, peripheral blood leukocyte subtypes were purchased from AlCells. Leukocytes were isolated from different, anonymous, nondiseased individuals’ whole blood (Supplementary Fig. S1) by magnetic-activated cell sorting (MACS) using a combination of negative and positive selection with highly specific cell surface antibodies conjugated to magnetic beads. The samples were obtained from men in 67%, those of white race in 41%, and with a mean age of 29 (SD = 9.0). The purity of separated cells was confirmed with flow cytometry to be ≥97% and included NK cells (n = 12), B cells (n = 5), T cells (n = 16), monocytes (n = 5), and granulocytes (n = 8). Genomic DNA was extracted and purified from cell pellets using a commercially available method (Qiagen), treated with sodium bisulfite (Zymo Research) and subjected to methylation profiling using the Infinium HumanMethylation27 BeadArray (Illumina).

This same platform was used for the analysis of samples from the case-control studies described later.

The HNSCC data set has been previously described (15) and consisted of 92 incident cases from the greater Boston area and 92 cancer-free population-based control subjects from the same region (18). The clinical characteristics for this study population are contained in Supplementary Table S2. The ovarian cancer data set (11) is publicly available from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/, Accession number GSE19711), and consisted of 266 postmenopausal women diagnosed with primary epithelial ovarian cancer (131 pretreatment and 135 posttreatment cases) from the UK Ovarian Cancer Population Study (UKOPS). Controls (n = 274) were cancer-free postmenopausal women for which annual serum samples were available. To avoid potential biases because of therapy, only pretreatment ovarian cases were included in our analysis. Clinical characteristics for the ovarian cancer data study population have been previously reported and can be found in Teschendorff and colleagues (11). The bladder cancer data set (13) consisted of 223 incident bladder cancer cases identified from the New Hampshire state cancer registry and 237 population controls from the same region (19, 20). Supplementary Table S3 provides a summary of the participant characteristics.

Statistical analysis
Our analytic strategy was aimed toward examining the extent to which peripheral blood DNA methylation of nonhematopoietic cancers is driven by the epigenetic signatures that define leukocyte subtypes. Linear mixed-effects models were used to assess differences in methylation across the leukocyte subtypes, modeling arcsine square root transformed methylation as the response for variance stabilization and normality considerations (21), leukocyte subtype as a fixed effect covariate, and a random effect term for plate/BeadChip. False discovery rate (FDR) estimation was used to control for the large number of comparisons and putative leukocyte DMRs were defined as those with FDR q-value < 0.05. Leukocyte DMRs were then ranked on the basis of the resulting q-values.

Methylation differences among the top 50 leukocyte DMRs were examined between cancer cases and cancer-free controls using a series of unconditional logistic regression models that were adjusted using available and relevant covariate information (see Fig. 1). A leukocyte DMR was considered differentially methylated if the nominal p-value from the unconditional logistic regression model was less than 0.05. Permutation tests were then applied to each of the 3 data sets to determine if the number of differentially methylated leukocyte DMRs was significantly greater than expected by chance. Specifically, samples were randomly permuted (same permutation across the top 50 DMRs) and an unconditional logistic regression model was fit to the resampled data. We considered 1,000 permutations for each data set to
generate the null distribution of the number of differentially methylated leukocyte DMRs. Permutation \( p \)-values were then obtained by comparing the observed number of differentially methylated leukocyte DMRs to the respective null distribution.

We next implemented an analysis that capitalized on the aggregate methylation signatures across a collection of leukocyte DMRs. Specifically for each cancer data set, we sought to train classifiers using the top \( M \) leukocyte DMRs followed by the validation of those classifiers in independent testing sets. This involved splitting each of the cancer data sets into equally sized training and testing sets, where the training sets were used to build the classifier and the respective testing set was used for the purposes of validation. Samples in the training set were clustered using the top \( M \) leukocyte DMRs, where \( M \) was determined for each training set from the total pool of putative DMRs using a previously described cross-validation procedure (22). We note that because the cross-validation procedure was implemented on each of the training data sets independently, there is no guarantee that the number \( M \) will be the same across the training sets. Clustering analysis was achieved using the Recursively Partitioned Mixture Model (RPMM, ref. 23), a hierarchical model-based method for clustering that has been extensively used for the clustering of array-based methylation data (13, 15, 24–26). On the basis of the RPMM fit to the training data, a naive Bayes classifier, a probabilistic classifier, was used to predict methylation class membership for the observations in the independent testing set. Associations between predicted methylation class and cancer case/control status were assessed using permutation \( \chi^2 \) tests and unconditional logistic regression models adjusted for available and relevant confounders. In addition, the classifier performance was investigated using receiver-operating characteristic (ROC) curves and the corresponding AUC.

We computed the pairwise Spearman correlation coefficients between the top \( M \) leukocyte DMRs and the CpG loci identified from the corresponding semisupervised RPMM (SS-RPMM; ref. 22) analysis of the HNSCC, ovarian, and bladder cancer data sets. A diagram illustrating the analytic framework for SS-RPMM is provided in Supplementary Fig. S1. Briefly SS-RPMM is a...
statistical methodology for identifying classes of methylation that are associated with a phenotype of interest and has been successfully applied in several of settings (13, 15, 27). We used the same training and testing sets for the previously described SS-RPMM analysis of the HNSCC and bladder cancer data sets (13, 15). This was done for the purposes of comparing the results of the present analysis to previously published results and to provide additional insight with respect to findings of those studies. For reasons of consistency, we also analyzed the ovarian cancer data set using the same SS-RPMM strategy and report those results in the Supplementary Data. Following the same logic as earlier, the same training and sets used for the SS-RPMM analysis were used for the leukocyte DMR profile analysis of the ovarian data. All analyses were carried out using the R statistical package, version 2.13 (www.r-project.org/).

Results

We began by profiling genome-wide DNA methylation in 46 samples of magnetic antibody sorted, normal human peripheral blood leukocyte subtypes (including B cells granulocytes, monocytes, NK cells, CD4+ T cells, CD8+ T cells, and Pan-T cells; Fig. 1A) using the Infinium HumanMethylation27 BeadArray. To discern leukocyte subtype DMRs, we examined the association between methylation and leukocyte subtype for each of the 26,486 autosomal CpG loci. This revealed 10,370 significantly differentially methylated CpGs among the leukocyte subtypes (fdr q-value < 0.05), which we ranked by q-value (Supplementary Table S4 and Fig. 1B). We selected the top 50 DMRs from this ranked list for use in the case-control analyses. Because the publically available ovarian cancer data set included both pre- and posttreatment cases, only pretreatment cases (n = 131) were considered in subsequent analyses to avoid potential biases resulting from therapy. Using unconditional logistic regression models, adjusted for available and relevant confounders (see Fig. 1), a substantial proportion of the 50 selected leukocyte DMRs were found to be significantly differentially methylated between cancer cases and cancer-free controls at the α = 0.05 threshold (46, 47, and 8 of 50, permutation P-values <0.001, <0.001, and 0.085, for HNSCC, ovarian cancer, and bladder cancer, respectively; Fig. 1C). For the ovarian data set, the largest difference between the β-values of controls and cancer cases among the proposed leukocyte DMRs was 11%, which is the largest difference in methylation between ovarian cases and controls considering all 26,486 autosomal CpGs. A similar finding was observed for the HNSCC data set, where the largest difference in methylation between cases and controls among leukocyte DMRs was 10%—which also corresponds to the largest difference in methylation between cases and controls across the array.

Of the leukocyte DMRs that were significantly differentially methylated in cancer cases compared with controls, 8 were common to all 3 cancer types (Fig. 1C). In HNSCC and ovarian cancer, 7 of these 8 leukocyte DMRs were hypomethylated in cases relative to controls, whereas all 8 DMRs were hypermethylated in bladder cancer cases relative to controls (Table 1).

To capitalize on the aggregate methylation signatures across a collection of leukocyte DMRs, we developed and tested classifiers based on profiles of leukocyte DMRs obtained from the subset analysis and subsequently assessed the performance of these classifiers for successfully discriminating cancer cases from cancer-free controls. Supplementary Figures S2 to S4 diagram the workflow of our DMR methylation profile analysis. For each of the 3 cancer data sets, a cross-validation procedure (22) was implemented on the training sets only to determine the number of top leukocyte DMRs (M) for subsequent clustering analysis of the training sets. On the basis of the respective cross-validation procedures using the 10,370 putative DMRs initially identified, the top 50, 10, and 56 leukocyte DMRs were selected to cluster the observations in the HNSCC, ovarian cancer, and bladder cancer

### Table 1. Methylation differences between cancer cases and controls for the 8 overlapping differentially methylated leukocyte DMRs

<table>
<thead>
<tr>
<th>Gene loci</th>
<th>HNSCC Mean Δβ (95% CI)</th>
<th>Ovarian Mean Δβ (95% CI)</th>
<th>Bladder Mean Δβ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C20orf135</td>
<td>−0.05 (−0.07 to −0.03)</td>
<td>−0.06 (−0.08 to −0.05)</td>
<td>0.02 (0.0−0.04)</td>
</tr>
<tr>
<td>PACAP</td>
<td>0.02 (0.00−0.04)</td>
<td>0.04 (0.02−0.05)</td>
<td>0.02 (0.0−0.04)</td>
</tr>
<tr>
<td>FGD2</td>
<td>−0.05 (−0.07 to −0.03)</td>
<td>−0.06 (−0.07 to −0.04)</td>
<td>0.02 (0.01−0.04)</td>
</tr>
<tr>
<td>SLC22A18</td>
<td>−0.05 (−0.07 to −0.04)</td>
<td>−0.05 (−0.06 to −0.04)</td>
<td>0.02 (0.01−0.04)</td>
</tr>
<tr>
<td>GSTP1</td>
<td>−0.05 (−0.07 to −0.04)</td>
<td>−0.06 (−0.07 to −0.05)</td>
<td>0.02 (0.01−0.04)</td>
</tr>
<tr>
<td>NFE2</td>
<td>−0.04 (−0.05 to −0.03)</td>
<td>−0.04 (−0.05 to −0.03)</td>
<td>0.02 (0.0−0.03)</td>
</tr>
<tr>
<td>ASGR2</td>
<td>−0.06 (−0.08 to −0.04)</td>
<td>−0.05 (−0.07 to −0.04)</td>
<td>0.02 (0.01−0.04)</td>
</tr>
<tr>
<td>SLC11A1</td>
<td>−0.05 (−0.07 to −0.04)</td>
<td>−0.05 (−0.04 to −0.06)</td>
<td>0.02 (0.0−0.04)</td>
</tr>
</tbody>
</table>

NOTE: Mean Δβ refers to the difference in mean methylation between cancer cases and controls (i.e., βcases − βcontrols).
training sets, respectively. The resultant clustering solutions were then used to predict methylation class membership for the subjects within the respective independent testing sets. Figures 2A, 3A, and 4A depict heat maps of the respective testing sets by predicted methylation class for each cancer data set. Methylation classes derived from leukocyte subtype DMRs were significantly associated with cancer case status within each cancer type.

Figure 2. Results from the DMR profile analysis of the HNSCC data set. A, heat map of the HNSCC testing data set. Rows represent subjects, which are grouped by predicted methylation class membership. Columns represent the top 50 leukocyte DMRs that were used to generate the methylation classes for the HNSCC testing set. Right, bar plot depicting the percent cancer case/control across the predicted methylation classes in the HNSCC testing set. B, ROC curves based on the predicted methylation classes only in the HNSCC testing set (blue) and methylation classes including patient age, gender, smoking status (never, former, current), smoking pack-years, weekly alcohol consumption, and HPV serostatus (orange).

Figure 3. Results from the DMR profile analysis of the ovarian data set. A, heat map of the ovarian testing data set. Rows represent subjects, which are grouped by predicted methylation class membership. Columns represent the top 10 leukocyte DMRs that were used to generate the methylation classes for the ovarian testing set. Right, bar plot depicting the percent cancer case/control across the predicted methylation classes in the ovarian testing set. B, ROC curves based on the predicted methylation classes alone in the ovarian testing set (blue) and methylation classes plus patient age group (55–60, 60–65, 65–70, 70–75, and >75 years; orange).
(permutation $\chi^2$ P-values $<0.0001$, $<0.0001$, and 0.03, HNSCC, ovarian cancer, and bladder cancer data sets, respectively), supporting the phenotypic relevance of predicted methylation classes based on leukocyte DMRs.

For the HNSCC testing set, subjects predicted to be in the right most classes of the dendrogram (classes beginning with R) were 6 times as likely to be HNSCC cases compared with subjects in the left most classes (classes beginning with L; OR = 5.99; 95% CI, 1.96–18.36), controlling for age, gender, smoking, alcohol consumption, and HPV serostatus. Assessing the classifier performance showed that methylation classes derived from the top 50 leukocyte DMRs were highly predictive HNSCC case/control status (AUC = 0.82; 95% CI, 0.74–0.91), which increased to 0.92 (0.87 with age, gender, smoking, and HPV serostatus included in the model; Fig. 2B). For ovarian cancer, subjects predicted to be in the right most classes were approximately 10 times as likely to be ovarian cancer cases compared with subjects in the left most classes (OR = 9.87; 95% CI, 4.63–21.10), controlling for age. In addition, the predicted methylation classes in the ovarian cancer data showed remarkably high sensitivity and specificity for predicting ovarian cancer case/control status (AUC = 0.82; 95% CI, 0.74–0.91), which increased to AUC = 0.86 (0.81 with age included in the model; Fig. 2B). In the bladder cancer data, subjects in the right most classes were nearly 2 times as likely to be cases compared with subjects in the left most (OR = 1.94; 95% CI, 0.95–3.98, adjusted for age, gender, smoking, and family history of bladder cancer), a somewhat less robust association than that observed for HNSCC and ovarian cancers. The classifier performance in the bladder cancer data was lower than that observed for HNSCC and ovarian cancer (bladder AUC = 0.67; 95% CI, 0.60–0.73 and adjusted AUC = 0.77; 95% CI, 0.71–0.83 with age, gender, smoking, and family history in the model; Fig. 4B).

Using leukocyte-derived DMRs to differentiate cases and controls results in methylation profiles that are consistent, and in the case of HNSCC and ovarian tumors, considerably better in terms of their prediction performance compared with previously published results using the same datasets (11, 13, 15). For the HNSCC and ovarian data sets, there was a high degree of correlation in the methylation status of leukocyte DMRs and CpG loci identified by previous analytic strategies (refs. 11, 15; mean absolute Spearman correlations = 0.68 and 0.75, respectively; Fig. 1). In contrast, the top 56 DMRs in the bladder data set were found to be less correlated with the CpG loci used to form the methylation classes in a previous study using the same data set (mean absolute Spearman correlation = 0.11; Fig. 5).

Discussion

Our novel investigation into the biologic underpinnings of disease-associated, blood-derived DNA methylation signatures in patients with solid tumors, suggest distinct,
Peripheral Blood DNA Methylation in Nonhematopoietic Cancers

Figure 5. Image plots representing the pairwise Spearman correlation coefficients. A, the 6 CpG loci identified by HNSCC analysis reported in ref. 15 and the top 50 leukocyte DMRs used in this analysis. B, the 7 CpG loci identified by the alternative ovarian analysis (Supplementary Fig. S5) and the top 10 leukocyte DMRs used in the present analysis. C, the 9 CpG loci identified by the bladder analysis reported in ref. 13 and the top 56 leukocyte DMRs used in this analysis.

well-defined immune-mediated responses to individual cancers. The motivation for our approach stems from the fact that blood-based assessments of DNA methylation are typically carried out using total WBC, therefore the methylation signatures responsible for distinguishing cancer cases and controls represent the aggregate methylation signatures across a complex cellular mixture of WBCs. As tumorigenesis elicits a distinct immune response (28–31), the result is a hematopoietic shift in WBC populations, which may be discerned by applying the unique epigenetic signature of differing lineages. Hence, the driving principle of this work is that the aggregate methylation signature in blood that distinguishes cancer cases from controls may in large part be because of the epigenetic signatures that define leukocyte subtypes. Given the chronic nature of cancer, it is possible that immune responses to comorbid conditions and treatment may be contributing to the methylation patterns reported here. We attempted to address the later by restricting our analysis to pretreatment cases, however such information was not available for the HNSCC and bladder data sets. It is also important to note that a prospective study with repeated DNA methylation assessments would be fundamental to conclusively determine whether altered methylation profiles occur as a result of an immune response to the tumor or are in some way promoting tumor growth and proliferation. At the same time, as a screening tool, even a methodology, which detects an early cancer, would hold clinical use.

To understand the role of immune-mediated responses to tumorigenesis in defining distinct signatures of blood-based DNA methylation between cancer cases and cancer-free controls, we studied the epigenetic landscape of WBCs by identifying DMRs among leukocyte subtypes. This analysis revealed that nearly all of the top 50 leukocyte DMRs were differentially methylated between cases and controls for HNSCC and ovarian cancers, with a smaller fraction differentially methylation between bladder cancer cases and controls. Among the 8 overlapping CpG loci that were found to be significantly differentially methylated between cancer cases and controls across the 3 data sets, the direction of the relationships was similar for HNSCC and ovarian cancer cases compared with controls opposed to that observed between bladder cancer cases and controls. These findings suggest that HNSCC and ovarian cancer may elicit similar shifts in leukocyte compositions in the hematopoietic system. Indeed, this finding is supported by recent work indicating an over-abundance of regulatory T cells (Tregs) in the tumor microenvironment of several types of cancer, including HNSCC and ovarian cancer (32–35). More specifically, it has been suggested that Tregs play a crucial role in the suppression of antitumor immune responses and thus participate in HNSCC progression and the immune escape process (36–39). Similarly, ovarian carcinoma cells are capable of producing TGF-β (40), a protein that regulates cellular proliferation and differentiation that is not only important for the functional integrity of Tregs, but also inhibits the proliferation and functional differentiation of T lymphocytes, NK cells, and macrophages (41, 42). Thus, it is possible that the tumor microenvironment may be contributing to the peripheral blood shifts we are observing using leukocyte DMRs.

Of the 8 overlapping DMRs (C20orf135, PACAP, FGD2, SLC22A18, GSTP1, NFE2, ASGR2, and SLC11A1) several are located within genes with either established or alleged involvement in immune differentiation or function (refs. 43–48; SLC11A1, PACAP, and FGD2). SLC11A1 is expressed in monocytes, the circulating precursors of dendritic cells, and macrophages (43, 44), which represent important antigen-presenting cells in the immune system. In addition, SLC11A1 has been shown to suppress IL-10 production (45), an antiinflammatory cytokine that strongly enhances B-cell survival and proliferation (46). Moreover, PACAP has been implicated as...
an intrinsic regulator of regulatory T-cell abundance after inflammation (47) and FGD2 has been shown to play a role in leukocyte signaling and vesicle trafficking in cells specialized to present antigen in the immune system (48).

Using our model containing the DNA methylation profile for the top 50 leukocyte DMRs, patient age, gender, smoking status, smoking pack-years, weekly alcohol consumption, and HPV serological status, HNSCC cancer was predicted with high degree of sensitivity and specificity. Similarly, high-prediction performance was obtained for ovarian cancer using the DNA methylation profile for the top 10 leukocyte DMRs and patient age group. Although the prediction performance for bladder cancer, based on the methylation profile of the top 56 DMRs, patient age, gender, smoking status, smoking pack-years, and family history of bladder cancer, was less than that observed for HNSCC and ovarian cancer, the AUC is consistent with a previous report (13). One explanation for the differences in magnitude for discriminating cancer cases and controls among cancer types is underlying differences in the magnitude of shift in leukocyte subtypes. Cancers characterized by a pronounced immunologic response such as HNSCC and ovarian cancer (49–53), may correspond to more discernable shifts in leukocyte subpopulations compared with bladder cancer (54), thus resulting in greater discrimination of blood-derived DNA methylation using leukocyte DMRs.

We also observed substantial correlation in methylation of the loci identified via the SS-RPMM analyses (ref. 15; Fig. 5 and Supplementary Fig. S5) and the leukocyte DMRs that defined the methylation classes discovered for the HNSCC and ovarian data sets. Given that the SS-RPMM procedure is specifically designed to construct methylation classes that are based on an optimal number of informative features (loci whose methylation is most strongly associated with cancer case/control status), our findings support the assertion that the methylation classes identified through SS-RPMM analyses of the HNSCC (15) and ovarian data sets (Supplementary Fig. S5) are in large part because of systematic hematopoietic changes in WBC populations in response to tumorigenesis. Contrary to the high correlation in methylation that was observed for the HNSCC and ovarian analyses, the 56 leukocyte DMRs used in the bladder profile analysis were notably less correlated with the 9 CpG loci identified via the previously reported SS-RPMM analysis of this data set (13). This may indicate a role for an alternative biologic mechanism in bladder cancer, where in addition to the epigenetic signatures characteristic of leukocyte subtypes, other epigenetic mechanisms independently contribute to the blood-derived differences in DNA methylation between bladder cancer cases and controls. Alternatively, it is also possible that our method for identifying leukocyte DMRs did not yield DMRs that are most important for bladder cancer immunobiology. More comprehensive profiling approaches across larger panels of leukocyte subpopulations are a high priority for future research.

Taken together, our results provide evidence that observed differences in blood-derived DNA methylation in cancer cases can be largely explained by systematic differences in the methylation signatures of leukocyte subpopulations. These findings signify that different cancers elicit a discernible immune response evident in peripheral blood. We believe these results have important implications for research into the immunology of cancer. Further, our approach provides a completely novel tool for the study of the immune profiles of diseases where only DNA can be accessed; that is, we believe this approach has use not only in cancer diagnostics and risk-prediction, but also can be applied to future research (including stored specimens) for any disease where the immune profile holds medical information. The approach described here is not capable of delineating the precise contributions of shifting leukocyte subpopulations to cancer-specific patterns in blood-based DNA methylation. However, work from our group has begun to address this issue (55).

In summary, our approach represents a simple, yet powerful and important new tool for medical research and may serve as a catalyst for future blood-based disease diagnostics.

Disclosure of Potential Conflicts of Interest
W. Accomando, E.A. Houseman, J.K. Wiencke, and K.T. Kelsey have ownership interest (including patents). No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C.J. Marsit, W. Accomando, M.R. Karagas, J.K. Wiencke, K.T. Kelsey
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): W. Accomando, M.R. Karagas, K.T. Kelsey
Study supervision: C.J. Marsit, M.R. Karagas, K.T. Kelsey

All authors discussed the results and commented on the manuscript.

Grant Support
This work was supported by the U.S. NIH grants (R01 CA121147, R01 CA107869, and R01 CA100679 to K.T. Kelsey; P42 ES007373 and R01 CA57494 to M.R. Karagas; R01 CA52689, NIEHS ES06717, P50 CA097257 to J.K. Wiencke) and the Flight Attendant Medical Research Institute (YCSA052541 to C.J. Marsit).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 11, 2012; revised May 31, 2012; accepted June 13, 2012; published OnlineFirst June 19, 2012.
References


Peripheral Blood Immune Cell Methylation Profiles Are Associated with Nonhematopoietic Cancers

Devin C. Koestler, Carmen J. Marsit, Brock C. Christensen, et al.