



Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2012 August ; 21(8): 1293–1302. doi:
10.1158/1055-9965.EPI-12-0361.

Peripheral blood immune cell methylation profiles are associated with non-hematopoietic cancers

Devin C. Koestler¹, Carmen J. Marsit^{1,2}, Brock C. Christensen^{1,2}, William Accomando⁴, Scott M. Langevin^{3,4}, E. Andres Houseman⁵, Heather H. Nelson⁶, Margaret R. Karagas¹, John K. Wiencke⁷, and Karl T. Kelsey^{3,4}

¹Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, Hanover, NH

²Department of Pharmacology and Toxicology, Geisel School of Medicine at Dartmouth, Hanover, NH

³Department of Pathology and Laboratory Medicine, Brown University, Providence, RI

⁴Department of Epidemiology, Brown University, Providence, RI

⁵Department of Public Health, Oregon State University, Corvallis, OR

⁶Division of Epidemiology and Masonic Cancer Center, University of Minnesota, Minneapolis, MN

⁷Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA

Abstract

Background—Blood leukocytes from patients with solid tumors exhibit complex and distinct cancer-associated patterns of DNA methylation. However, the biological mechanisms underlying these patterns remain poorly understood. Since 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 (DMRs) 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 to cancer-free controls (48 and 47 out of 50, respectively). Methylation classes derived from leukocyte DMRs were significantly associated 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 (AUC = 0.82, 0.83, and 0.67).

Corresponding Author: Karl T. Kelsey, MD, MOH, Department of Pathology and Laboratory Medicine, 70 Ship St. Brown University, Providence, RI 02912, Phone: 401-863-6420, Karl_Kelsey@brown.edu.

Competing financial interests

The authors declare no competing financial interests.

Author contributions

D.C.K. implemented the statistical analyses and prepared the manuscript with C.J.M., B.C.C., S.M.L., H.H.N., M.R.K., J.K.W., and K.T.K. E.A.H. provided statistical guidance and analytic support and W.A., contributed to the analysis and discussion. All authors discussed the results and commented on the manuscript

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

Impact—This illustrates the potential utility 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.

Keywords

DNA methylation; cancer; leukocytes; immune system; biomarkers

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 non-hematopoietic 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, 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) (15), and in a two 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 cancer patients could offer important new insights into the pathophysiology of cancer while also serving as a promising new avenue for non-invasive cancer detection and diagnostics. Despite these highly significant findings however, the biological 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 over-represented 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.

Due to the potential clinical utility 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 (DMRs) that differentiate leukocyte subtypes. These leukocyte-tagging DMRs were then investigated using epigenome-wide blood methylation data from three

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.

Methods

Study Population

Sorted, normal, human, peripheral blood leukocyte subtypes were purchased from AllCells (Emeryville, CA). Leukocytes were isolated from different, anonymous, non-diseased individuals' whole blood (Supplementary Figure 1) 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, Valencia, CA), treated with sodium bisulfite (Zymo Research, Irvine, CA) and subjected to methylation profiling using the Infinium HumanMethylation27 BeadArray (Illumina, San Diego, CA). This same platform was used for the analysis of samples from the case-control studies described below.

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 2. 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 pre-treatment and 135 post-treatment 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 due to therapy, only pre-treatment 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 et al. (2009). 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 3 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 non-hematopoietic 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 false discovery rate (fdr) q-value < 0.05. Leukocyte DMRs were then ranked based on 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 Figure 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 three 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 1000 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 since 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 (23) (RPMM), a hierarchical model-based method for clustering that has been extensively used for the clustering of array-based methylation data (13,15, 24–26). Based on 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 χ^2 tests and unconditional logistic regression models adjusted for available and relevant confounders. Additionally, the classifier performance was investigated using receiver operating characteristic (ROC) curves and the corresponding area under the curve (AUC).

We computed the pairwise spearman correlation coefficients between the top M leukocyte DMRs and the CpG loci identified from the corresponding semi-supervised RPMM (22) (SS-RPMM) analysis of the HNSCC, ovarian, and bladder cancer data sets. A diagram illustrating the analytic framework for SS-RPMM is provided in Supplementary Figure 1. 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 above, 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; Figure 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 4 and Figure 1b). We selected the top 50 DMRs from this ranked list for use in the case-control analyses. Since the publically available ovarian cancer data set included both pre- and post-treatment cases, only pre-treatment 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 Figure 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 $\mu = 0.05$ threshold (48, 47, and 8 out of 50, permutation p-values = <0.001, <0.001, 0.085, for HNSCC, ovarian cancer, and bladder cancer, respectively) (Figure 1c). For the ovarian data set, the largest difference between the beta-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 26486 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 to controls, eight were common to all three cancer types (Figure 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 2–4 diagram the workflow of our DMR methylation profile analysis. For each of the three 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. Based on 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 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,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 (permutation χ^2 p-values <0.0001, <0.0001, 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 to 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 demonstrated that methylation classes derived from the top 50 leukocyte DMRs were highly predictive HNSCC case/control status (area under the curve (AUC) = 0.82 95% CI [0.74, 0.91]), which increased to 0.92 (0.87, 0.98 with age, gender, smoking, alcohol consumption, and HPV serostatus included in the model (Figure 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 to subjects in the left most classes (OR = 9.87, 95% CI [4.63, 21.10]), controlling for age. Additionally, the predicted methylation classes in the ovarian cancer data demonstrated remarkably high sensitivity and specificity for predicting ovarian cancer case/control status (AUC = 0.83 95% CI [0.77, 0.89]), which increased to AUC = 0.86 95% CI [0.81, 0.92] with age included in the model (Figure 3b). In the bladder cancer data, subjects in the right most classes were nearly 2 times as likely to be cases compared to 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) (Figure 4b).

Utilizing 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 to 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 (11,15) (mean absolute spearman correlations = 0.68 and 0.75, respectively; Figure 5). 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; Figure 5).

Discussion

Our novel investigation into the biological underpinnings of disease-associated, blood-derived DNA methylation signatures in patients with solid tumors, suggest distinct, 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 due to the epigenetic signatures that define leukocyte subtypes. Given the chronic nature of cancer, it is possible that immune responses to co-morbid conditions and treatment may be contributing to the methylation patterns reported here. We attempted to address the later by restricting our analysis to pre-treatment 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 utility.

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 methylated between bladder cancer cases and controls. Among the eight overlapping CpG loci that were found to be significantly differentially methylated between cancer cases and controls across the three data sets, the direction of the relationships was similar for HNSCC and ovarian cancer cases compared to 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 overabundance 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 anti-tumor 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 eight overlapping DMRs (*C20orf135*, *PACAP*, *FGD2*, *SLC22A18*, *GSTP1*, *NFE2ASGR2*, and *SLC11A1*) several are located within genes with either established or alleged involvement in immune differentiation or function (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. Additionally, *SLC11A1* has been shown to suppress IL-10 production (45), an anti-inflammatory 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 sub-populations compared to 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 (15) (Figure 5 and Supplementary Figure 5) 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 Figure 5) are in large part due to 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 biological 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 leukocytes subpopulations is 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 sub-populations. 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 utility 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This work was supported by the US National Institutes of Health grants (R01 CA121147, R01 CA078609, and R01 CA100679 to K.T.K., P42 ES007373 and R01 CA57494 to M.R.K., R01 CA52689, NIEHS ES06717, P50 CA097257 to J.K.W.); and the Flight Attendant Medical Research Institute (YCSA052341 to C.J.M.)

References

1. Shen L, Kondo Y, Guo Y, Zhang J, Zhang L, Ahmed S, et al. Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. *PLoS Genetics*. 2007; 3(10):e181. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2041996/>.
2. Sincic N, Herceg Z. DNA methylation and cancer: ghosts and angels above the genes. *Curr Opin Oncol*. 2011; 23(1):23, 69–76.

3. Cheung HH, Lee TL, Rennert OM, Chan WY. DNA Methylation of Cancer Genome. *Birth Defects Res C*. 2009; 87(4):335–50.
4. Cui HM. Loss of imprinting of IGF2 as an epigenetic marker for the risk of human cancer. *Dis Markers*. 2007; 23(1–2):105–112. [PubMed: 17325430]
5. Wilhelm-Benartzi CS, Koestler DC, Houseman EA, Christensen BC, Wiencke JK, Schend AR, et al. DNA methylation profiles delineate etiologic heterogeneity and clinically important subgroups of bladder cancer. *Carcinogenesis*. 2010; 31(11):1972–1976. [PubMed: 20802236]
6. Schwartzman J, Mongoue-Tchokote S, Gibbs A, Gao L, Corless CL, Jin J, et al. A DNA methylation microarray-based study identifies ERG as a gene commonly methylated in prostate cancer. *Epigenetics*. 2011; 6(10):1248–1256. [PubMed: 21946329]
7. Christensen BC, Houseman EA, Godleski JJ, Marsit CJ, Longacker JL, Roelofs CR, et al. Epigenetic profiles distinguish pleural mesothelioma from normal pleura and predict lung asbestos burden and clinical outcome. *Cancer Res*. 2009; 69(1):227–234. [PubMed: 19118007]
8. Marsit CJ, Houseman EA, Christensen BC, Eddy K, Bueno R, Sugarbaker DJ, et al. Examination of a CpG island methylator phenotype and implications of methylation profiles in solid tumors. *Cancer Res*. 2006; 66(21):10621–10629. [PubMed: 17079487]
9. Marsit CJ, Christensen BC, Houseman EA, Karagas MR, Wrench MR, Yeh RF, et al. Epigenetic profiling reveals etiologically distinct patterns of DNA methylation in head and neck squamous cell carcinoma. *Carcinogenesis*. 2009; 30(3):416–422. [PubMed: 19126652]
10. Widschwendter M, Apostolidou S, Raum E, Rothenbacher D, Fiegl H, Menon U, et al. Epigenotyping in peripheral blood cell DNA and breast cancer risk: a proof of principle study. *PLoS One*. 2008; 3:e2656. [PubMed: 18628976]
11. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Gayther SA, Apostolidou S, et al. An epigenetic signature in peripheral blood predicts active ovarian cancer. *PLoS One*. 2009; 3(7):e8274. <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0002656>. [PubMed: 20019873]
12. Wang L, Aakre JA, Jiang R, Marks RS, Wu Y, Chen J, et al. Methylation markers for small cell lung cancer in peripheral blood leukocyte DNA. *J Thorac Oncol*. 2010; 5(6):778–785. [PubMed: 20421821]
13. Marsit CJ, Koestler DC, Christensen BC, Karagas MR, Houseman EA, Kelsey KT. DNA methylation array analysis identifies profiles of blood-derived DNA methylation associated with bladder cancer. *J Clin Oncol*. 2011; 29(9):1133–1139. [PubMed: 21343564]
14. Pedersen KS, Bamlet WR, Oberg AL, de Andrade M, Matsumoto ME, Tang H, et al. Leukocyte DNA methylation signature differentiates pancreatic cancer patients from healthy controls. *PLoS One*. 2011; 6(3):e18223. <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0018223>. [PubMed: 21455317]
15. Langevin SM, Koestler DC, Christensen BC, Butler RA, Wiencke JK, Nelson HH, et al. Peripheral blood DNA methylation profiles are predictive of head and neck squamous cell carcinoma: an epigenome-wide association study. *Epigenetics*. 2012 In press.
16. Laird PW. The power and the promise of DNA methylation markers. *Nat Rev Cancer*. 2003; 3(4):253–266. [PubMed: 12671664]
17. Laird PW. Cancer epigenetics. *Hum Mol Genet*. 2005; 14(1):65–76.
18. Applebaum KM, McClean MD, Nelson HH, Marsit CJ, Christensen BC, Kelsey KT. Smoking modifies the relationship between XRCC1 haplotypes and HPV6-negative head and neck squamous cell carcinoma. *Int J Cancer*. 2009; 124(11):2690–2696. [PubMed: 19230024]
19. Karagas MR, Tosteson TD, Blum J, Morris JS, Baron JA, Klaue B. Design of an epidemiologic study of drinking water arsenic exposure and skin and bladder cancer risk in a US population. *Environ Health Perspect*. 1998; 106(4):1047–1050. [PubMed: 9703491]
20. Wallace K, Kelsey KT, Schned A, Morris JS, Andrew AS, Karagas MR. Selenium and risk of bladder cancer: a population-based case-control study. *Cancer Prev Res*. 2009; 2(1):70–73.
21. Rocke DM. On the beta transformation family. *Technometrics*. 1993; 35(1):72–81.
22. Koestler DC, Marsit CJ, Christensen BC, Karagas MR, Bueno R, Sugarbaker DJ, et al. Semi-supervised recursively partitioned mixture models for identifying cancer subtypes. *Bioinformatics*. 2010; 26(20):2578–2585. [PubMed: 20834038]

23. Houseman EA, Christensen BC, Yeh RF, Marsit CJ, Karagas MR, Wrench MR, et al. Model-based clustering of DNA methylation array data: a recursive-partitioning algorithm for high-dimensional data arising a mixture of beta distributions. *BMC Bioinformatics*. 2008; 9:365. <http://www.biomedcentral.com/1471-2105/9/365>. [PubMed: 18782434]
24. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrench MR, Wiemels JL, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet*. 2009; 5(8):e1000602. <http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1000602>. [PubMed: 19680444]
25. Christensen BC, Smith AA, Zheng S, Koestler DC, Houseman EA, Marsit CJ, et al. DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma. *J Natl Cancer Inst*. 2011; 103(2):143–153. [PubMed: 21163902]
26. Hinoue T, Weisenberger DJ, Lange CP, Shen H, Byun HM, Van Den Berg D, et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer [published online ahead of print June 9, 2011]. *Genome Res*. 2011;10.1101/gr.117523.11
27. Banister CE, Koestler DC, Maccani MA, Padbury JF, Houseman EA, Marsit CJ. Infant growth restriction is associated with distinct patterns of DNA methylation in human placentas. *Epigenetics*. 2011; 6(7):920–927. [PubMed: 21758004]
28. Yamanaka T, Matsumoto S, Teramukai S, Ishiwata R, Nagai Y, Fukushima M. The baseline ratio of neutrophils to lymphocytes is associated with patient prognosis in advanced gastric cancer. *Oncology*. 2007; 73(3–4):215–220. [PubMed: 18424885]
29. Ji H, Houghton AM, Mariani TJ, Perera S, Kim CB, Padera R, et al. K-ras activation generates an inflammatory response in lung tumors. *Oncogene*. 2006; 25(14):2105–2112. [PubMed: 16288213]
30. Rui L, Schmitz R, Ceribelli M, Staudt LM. Malignant pirates of the immune system. *Nat Immunol*. 2011; 12:933–940. [PubMed: 21934679]
31. Whiteside TL. Immune responses to malignancies. *J Allergy Clin Immunol*. 2010; 125(2):272–283.
32. Alhamarneh O, Amarnath SMP, Stafford ND, Greenman J. Regulatory T cells: what role do they play in antitumor immunity in patients with head and neck cancer. *Head Neck*. 2008; 30(2):251–261. [PubMed: 18172882]
33. Schott AK, Pries R, Wollenberg B. Permanent up-regulation of regulatory T-lymphocytes in patients with head and neck cancer. *Int J Mol Med*. 2010; 26(1):67–75. [PubMed: 20514424]
34. Peng DJ, Liu R, Zou W. Regulatory T-cells in ovarian cancer. *J Oncol*. 2012:345164. [PubMed: 22481922]
35. Alevaro AB, Montagna MK, Craveiro V, Liu L, Mor G. Distinct subpopulations of epithelial ovarian cancer cells can differentially induce macrophages and T regulatory cells toward a pro-tumor phenotype. *Am J Reprod Immunol*. 2012; 67(3):256–65. [PubMed: 21917055]
36. Whiteside TL. Immunobiology of head and neck cancer. *Cancer Metastasis Rev*. 2005; 24:95–105. [PubMed: 15785875]
37. McHugh RS, Shevach EM. The role of suppressor T cells in regulation of immune responses. *J Allergy Clin Immunol*. 2002; 110:693–702. [PubMed: 12417876]
38. Szczepanski MJ, Szajnik M, Czystowska M, Mandapathil M, Strauss L, Welsh A, et al. Increased frequency and suppression by regulatory T cells in patients with acute myelogenous leukemia. *Clin Cancer Res*. 2009; 15:3325–3332. [PubMed: 19417016]
39. Whiteside TL. Immunobiology and immunotherapy of head and neck cancer. *Curr Oncol Rep*. 2001; 3:46–55. [PubMed: 11123869]
40. Bartlett JM, Langdon SP, Scott WN, Love SB, Miller EP, Katsaros D, et al. Transforming growth factor- β isoform expression in human ovarian tumours. *European Journal of Cancer*. 1997; 33(14):2397–2403. [PubMed: 9616289]
41. Chen W, Jin W, Hardegen N, Lei K, Li L, Marinos N, et al. Conversion of peripheral CD4⁺CD25⁻ naive T cells to CD4⁺CD25⁺ regulatory T cells by TGF- β induction of transcription factor Foxp3. *Journal of Experimental Medicine*. 2003; 198(12):1875–1886. [PubMed: 14676299]
42. Cazac BB, Roes J. TGF- β receptor controls B cell responsiveness and induction of IgA in vivo. *Immunity*. 2000; 13(4):443–451. [PubMed: 11070163]
43. BioGPS. [<http://biogps.gnf.org>]

44. HaemAtlas. [<http://dil.t1dbase.org/page/HaemAtlasView>]
45. Fritsche G, Nairz M, Werner ER, Barton HC, Weiss G. Nramp1-functionality increases iNOS expression via repression of IL-10 formation. *Eur J Immunol.* 2008; 38(11):3060–3067. [PubMed: 18991287]
46. Sabat R, Grutz G, Warszawska K, Kirsch S, Witte E, Wolk K, et al. Biology of interleukin-10. *Cytokine Growth Factor.* 2010; 21:331–344.
47. Tan YV, Abad C, Lopez R, Dong H, Liu S, Lee A, et al. Pituitary adenylyl cyclase-activating polypeptide is an intrinsic regulator of treg abundance and protects against experimental autoimmune encephalomyelitis. *Proc Natl Acad Sci.* 2009; 106(6):2012–2017. [PubMed: 19190179]
48. Huber C, Mrtensson A, Bokoch GM, Nemazee D, Gavin AL. Fgd2, a cdc42-specific exchange factor expressed by antigen-presenting cells, localizes to early endosomes and active membrane ruffles. *J Biol Chem.* 2008; 283(49):34002–34012. [PubMed: 18838382]
49. Tong CC, Kao J, Sikora AG. Recognizing and reversing the immunosuppressive tumor microenvironment of head and neck cancer. *Immunol Res.* 2012 [In press].
50. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med.* 2003; 348(3):203–213. [PubMed: 12529460]
51. Tomsova M, Melichar B, Sedlakova I, Steiner I. Prognostic significance of CD3+ tumor-infiltrating lymphocytes in ovarian carcinoma. *Gynecol Oncol.* 2008; 108(2):415–420. [PubMed: 18037158]
52. Sato E, Olson SH, Ahn J, Bundy B, Nishikawa H, Qian F, et al. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. *Proc Natl Acad Sci.* 2005; 102(51):18538–18543. [PubMed: 16344461]
53. Curiel TJ, Coukos G, Zou L, Alvarez X, Cheng P, Mottram P, et al. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nat Med.* 2004; 10(9):942–949. [PubMed: 15322536]
54. Soygur T, Beduk Y, Yaman O, Yilmaz E, Tokgoz G, Gogus O. Analysis of the peripheral blood lymphocyte subsets in patients with bladder cancer. *Urology.* 1999; 53(1):88–91. [PubMed: 9886594]
55. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012; 13(1):86. [PubMed: 22568884]

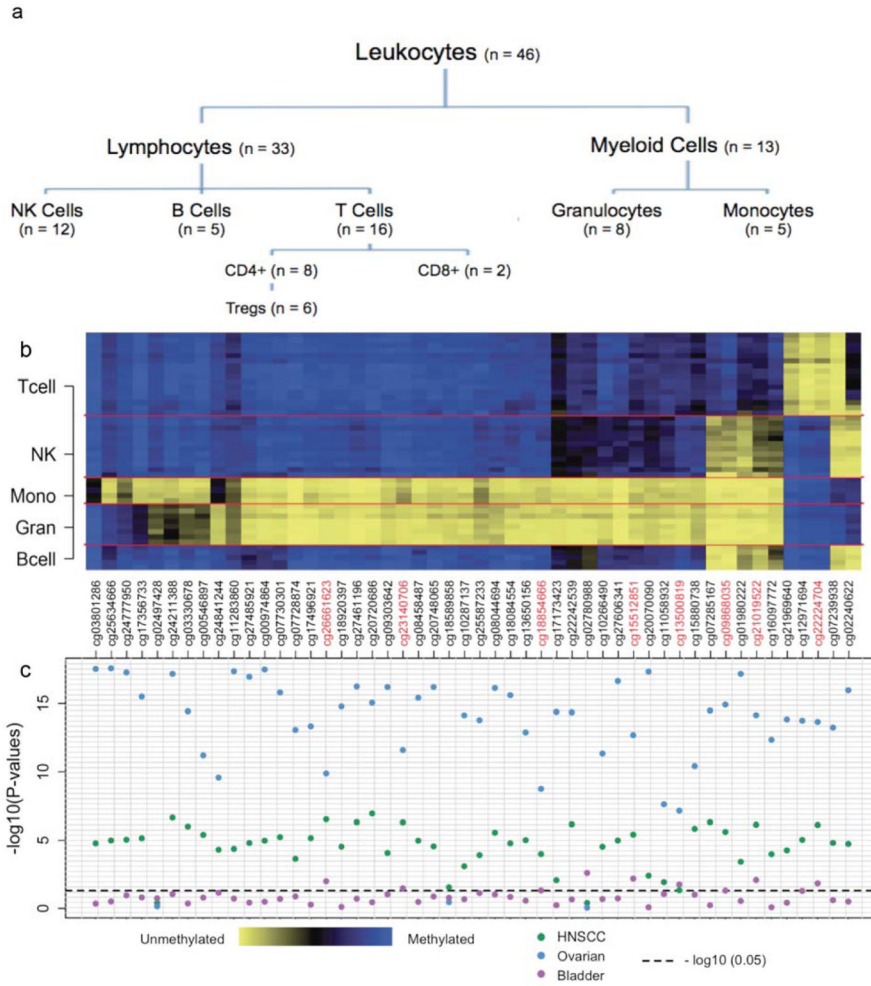


Figure 1. Results from the DMR subset analysis

(a) Hierarchy for the leukocyte subtypes. (b) Heat map of the methylation status for the top 50 leukocyte DMRs by leukocyte subtype. (c) Plot depicting the $-\log_{10}(\text{p-values})$ for the top 50 leukocyte DMRs across the three cancer data sets (blue - HNSCC; green - Ovarian; purple - Bladder;). P-values capture methylation differences between cancer cases and non-cancer controls and were obtained from individual unconditional logistic regression models fit to each of the 50 leukocyte DMRs. For the HNSCC data set, logistic regression models were adjusted for patient age, gender, smoking status (never, former, current), smoking pack years, weekly alcohol consumption, and HPV serology status. The bladder cancer data set was adjusted for patient age, gender, smoking status, smoking pack years, and family history of bladder cancer and the ovarian cancer data set was adjusted for patient age group (55–60, 60–65, 65–70, 70–75 and >75 years). The horizontal dashed line represents $-\log_{10}(p = 0.05)$.

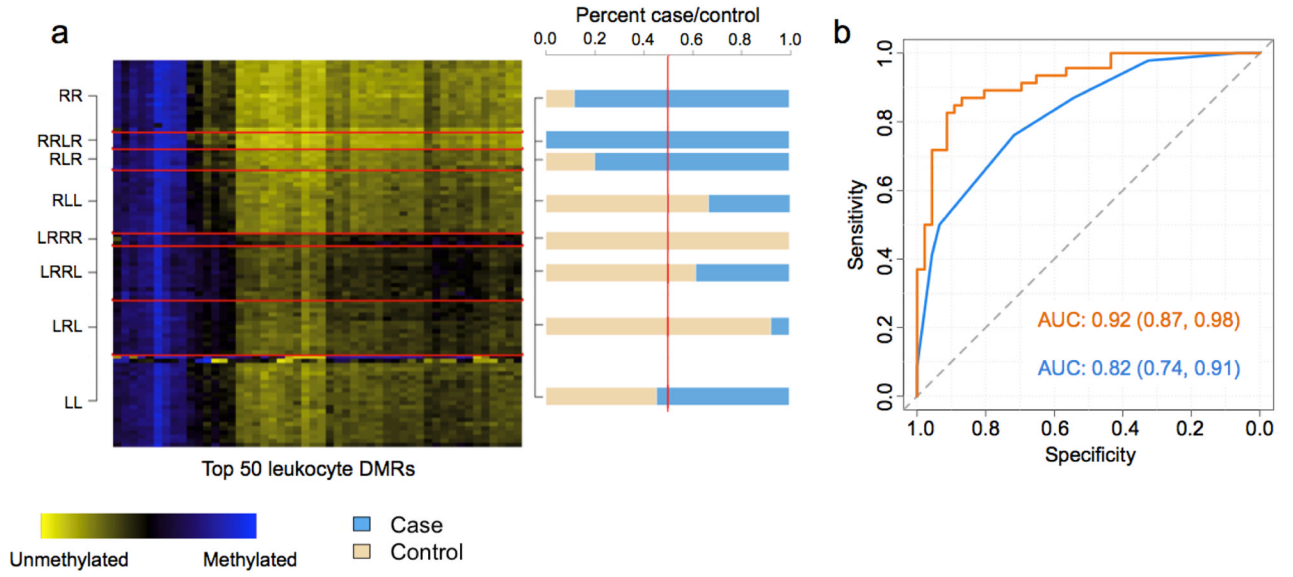


Figure 2. Figures depicting the results from the DMR profile analysis of the HNSCC data set (a) Heat map of the HNSCC testing data set. Rows represent subjects, which are grouped by predicted methylation class membership. Columns represent the top 50 leukocyte DMRs that were used to generate the methylation classes for the HNSCC testing set. The right column of (a) represents a 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).

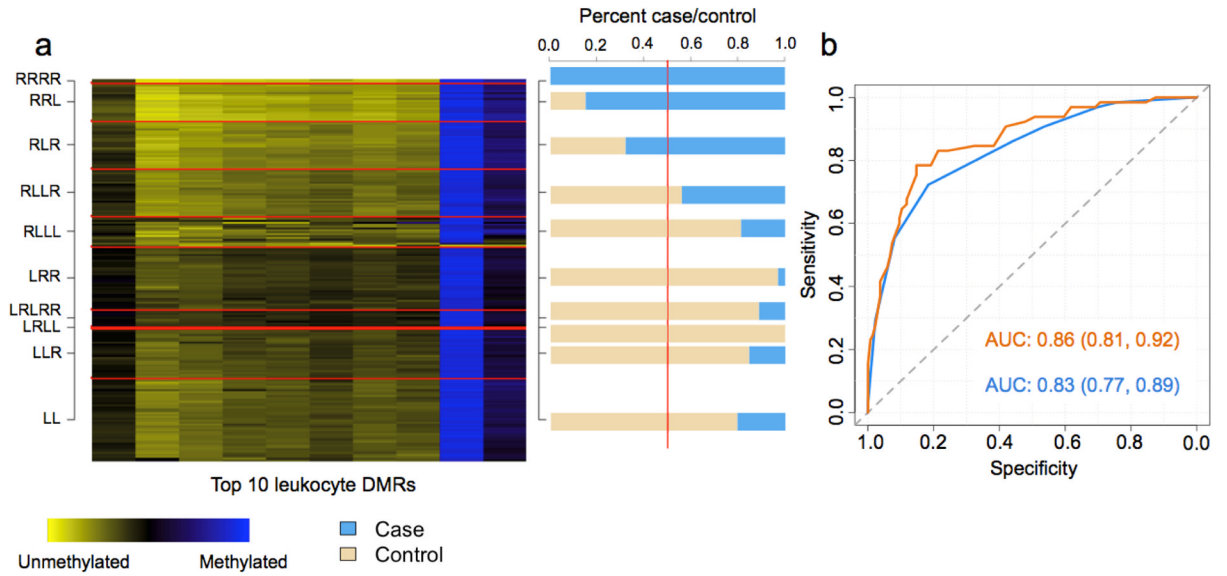


Figure 3. Figures depicting the results from the DMR profile analysis of the Ovarian data set (a) Heat map of the ovarian testing data set. Rows represent subjects, which are grouped by predicted methylation class membership. Columns represent the top 10 leukocyte DMRs that were used to generate the methylation classes for the ovarian testing set. The right column of (a) represents a 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).

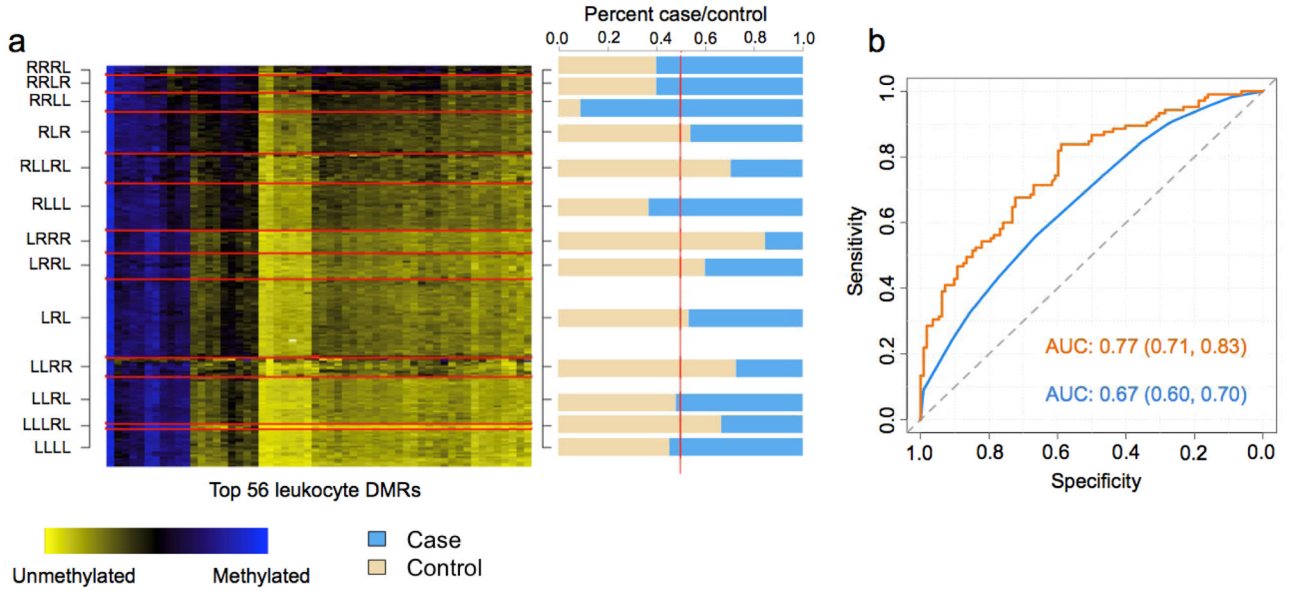


Figure 4. Figures depicting the results from the DMR profile analysis of the bladder data set (a) Heat map of the bladder testing data set. Rows represent subjects, which are grouped by predicted methylation class membership. Columns represent the top 56 leukocyte DMRs that were used to generate the methylation classes for the bladder testing set. The right column of (a) represents a bar-plot depicting the percent cancer case/control across the predicted methylation classes in the bladder testing set. (b) ROC curves based on the predicted methylation classes alone in the bladder testing set (blue) and methylation classes plus patient age, gender, smoking status (never, former, current), smoking pack years, and family history of bladder cancer (orange).

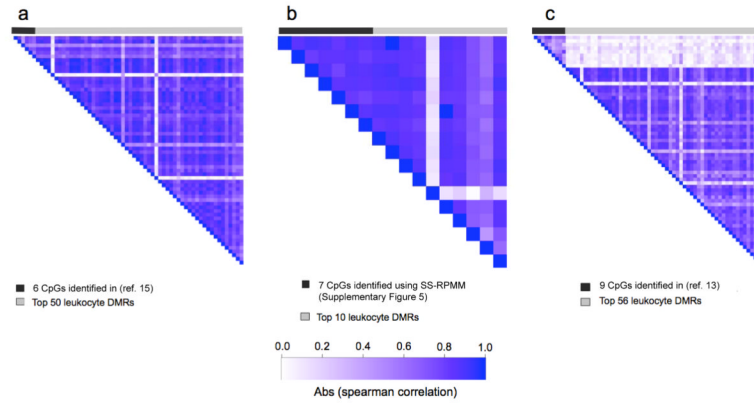


Figure 5. Image plots representing the pairwise spearman correlation coefficients (a) The 6 CpG loci identified by HNSCC analysis reported in (ref. 15) and the top 50 leukocyte DMRs used in the present analysis, (b) the 7 CpG loci identified by the alternative ovarian analysis (Supplementary Figure 5) and the top 10 leukocyte DMRs used in the present analysis, and (c) the 9 CpG loci identified by the bladder analysis reported in (ref. 13) and the top 56 leukocyte DMRs used in the present analysis.

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

Mean delta-beta refers to the difference in mean methylation between cancer cases and controls (i.e. $\beta_{\text{cases}} - \beta_{\text{controls}}$).

Gene Loci	Mean delta-beta (95% CI)		
	HNSCC	Ovarian	Bladder
C20orf135	-0.05 (-0.07, -0.03)	-0.06 (-0.08, -0.05)	0.02 (0.0, 0.04)
PACAP	0.02 (0.00, 0.04)	0.04 (0.02, 0.05)	0.02 (0.0, 0.04)
FGD2	-0.05 (-0.07, -0.03)	-0.06 (-0.07, -0.04)	0.02 (0.01, 0.04)
SLC22A18	-0.05 (-0.07, -0.04)	-0.05 (-0.06, -0.04)	0.02 (0.01, 0.04)
GSTP1	-0.05 (-0.07, -0.04)	-0.06 (-0.07, -0.05)	0.02 (0.01, 0.04)
NFE2	-0.04 (-0.05, -0.03)	-0.04 (-0.05, -0.03)	0.02 (0.0, 0.03)
ASGR2	-0.06 (-0.08, -0.04)	-0.05 (-0.07, -0.04)	0.02 (0.01, 0.04)
SLC11A1	-0.05 (-0.07, -0.04)	-0.05 (-0.04, -0.06)	0.02 (0.0, 0.04)