Reversal DNA methylation patterns for cancer diagnosis

Subtitle: DNA methylation for cancer diagnosis

Hongdong Li, Guini Hong, Zheng Guo
Bioinformatics Centre,
School of Life Science, University of Electronic Science and Technology of China
Chengdu, China
email: biomantis_lhd@163.com, hongguini08@gmail.com, guoz@ems.hrbmu.edu.cn

Zheng Guo
Department of Bioinformatics,
School of Basic Medical College, Fujian Medical University,
Fuzhou, China
email: guoz@ems.hrbmu.edu.cn

***Abstract*—Detecting aberrant DNA methylation as diagnostic or prognostic biomarkers for cancer has been a topic of considerable interest recently. However, current classifiers based on absolute methylation values detected from a cohort of samples are typically difficult to be transferable to other cohorts of samples. Here, we employed a modified rank-based method to extract pairs of CpG sites with reversal relative DNA methylation levels in disease samples to those in normal controls for five cancer types respectively. The reversal pairs showed excellent prediction performance with the accuracy above 95% for each type of cancer. Furthermore, the reversal pairs identified for a cancer type could distinguish samples with different subtypes and different malignant stages including early stage of this cancer from normal controls and were also specific to this cancer. In conclusion, the reversal pairs detected by the rank-based method are accurate and transferable to independent cohorts of samples, which are also applicable to early cancer diagnosis. They could also be used to detect common molecular alterations in cancer, which can shed light on the other follow-up studies.**

Keywords—DNA methylation; cancer; rank-based;

I. INTRODUCTION

Aberrant DNA methylation in cancer, including global hypomethylation and local hypermethylation of certain genes, is one of the common forms of molecular alterations in carcinogenesis [1]. It has been recognized that DNA-based molecular biomarkers, such as DNA methylation, are readily amplifiable and easily transferable from a research laboratory setting into routine diagnostics in a clinical trials [2]. Therefore, many researchers have tried to detect aberrant DNA methylation as diagnostic or prognostic biomarkers for various types of cancer in the past few years [2-4]. However, the transferability of many detected biomarkers usually lack of validation in independent datasets. On the other hand, classifiers constructed by using machine learning methods, such as support vector machine (SVM) [5] and artificial neural networks [6], are difficult to interpret biological meaning according to the rules of classification and hardly transferable to independent experiments. Thus, it is necessary to find a way to extract the methylation patterns with readily interpretable rules that can accurately discriminate tumor and normal samples from methylation profiles. The relative expression-based method for finding disease biomarkers entirely based on pairs of genes with relative expression values in disease samples reversal to those in the controls. Comparing with other machine learning methods, this parameter-free method can avoid data over-fitting and classifiers obtained by this method are biologically interpretable, transferable, and invariable to any monotonic transformation of the data [7-8]. It has been used in array-based gene expression successfully. However, whether this relative ordering of genes can be applied to analyze DNA methylation data has not yet been evaluated.

Thus, in this study, based on DNA methylation profiles collected from TCGA (the Cancer Genome Atlas, http://tcga-data.nci.nih.gov/tcga) database, we employed the rank-based methods to detect the relative methylation level reversal pairs (R-pairs) between normal and tumor tissue samples as candidate marker pairs. Then, we identified the most discriminatory R-pairs by considering the top CpG sites with highest appearance frequencies in all candidate marker pairs. For simplicity, we only focused on the top 11 R-pairs (FR-pairs) involved 11 CpG sites with the highest appearance frequencies in all candidate marker pairs. The FR-pairs identified for each cancer type performed well in testing sets for the same cancer types and also in validation sets from other independent experiments. Moreover, they also performed well in distinguishing samples with different subtypes and different malignant degrees for the same cancer types from normal controls. Validation on DNA methylation profiles from different cancer types showed that the FR-pairs were specific to cancer type.

II. MATERIAL AND METHODS

A. Datasets

The DNA methylation profiles analyzed in this study were downloaded from the Gene Expression Omnibus (GEO) [9] and TCGA database. Detailed dataset information was described in Table I. All profiles were generated using the Human Methylation27 Bead Array (San Diego, CA, USA), targeting 27578 CpG sites located in promoter regions of unique 14495 genes. For the datasets collected from TCGA,

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

TABLE I. METHYLATION DATASETS USED IN THIS STUDY

Cancer Type	Abbreviation	Sample Size		Data Source
		Normal	*Tumor*	
Lung adenocarcinoma	LUAD	24	127	TCGA
Kidney renal clear cell carcinom	KIRC	199	219	TCGA
Colon adenocarcinoma	COAD	37	167	TCGA
Stomach adenocarcinoma	STAD	57	80	TCGA
Breast invasive carcinoma	BRCA	27	315	TCGA
Lung squamous cell carcinoma	LUSC	27	133	TCGA
Kidney renal papillary cell carcinoma	KIRP	5	16	TCGA
colorectal cancer	CRC44	22	22	GEO(GSE17648)
gastric cancer	GAC75	32	43	GEO(GSE25869)
Breast cancer	BRC248	12	236	GEO(GSE20713)

only level 2 data were used, which included methylated signal intensity (M) and unmethylated signal intensity (U) for each probe. For each CpG site, the methylation level, denoted as a beta-value (β), was calculated as below [10]:

$$\beta = \frac{\max(M,0)}{U + M + 100} \tag{1}$$

Each of the five cancer datasets, namely lung adenocarcinoma (LUAD), kidney renal clear cell carcinoma (KIRC), colon adenocarcinoma (COAD), stomach adenocarcinoma (STAD) and breast invasive carcinoma (BRCA), was divided into two subsets according to the batch ID provided by the TCGA database: the batch comprising of normal and cancer samples with the largest normal and cancer sample sizes as training set and the remaining batches as testing set (Table SI). Apart from these five cancer datasets, the remaining datasets used as validation sets.

B. Detection of R-pairs

For each training dataset for each of the five cancer types, we determined the relative methylation level reversal pair (R-pair) based on the TSP method [7]. For a given dataset, the methylation profiles can be represented as a matrix *A* with dimension $M \times N$, where *M* represents the number of CpG sites and *N* represents the number of profiles. A profile either belongs to *class1* (normal samples) or *class2* (tumor samples) and could be denoted as $[\beta_1,...,\beta_i,...,\beta_M]$, where β_i represent the methylation level for CpG site *i*. If the methylation levels of two CpG sites, *k* and *j*, satisfied that the probability of $\beta_k < \beta_j$ in *class1* significantly differed from that in *class2*, these two CpG sites can be considered as R-pair.

Suppose there are *N1* samples in *class1* and *N2* samples in *class2* (*N1+N2=N*). For a R-pair (*k, j*), if $\beta_k < \beta_j$ was observed in *a* samples in *class1* and *b* samples in *class2*, then the difference in probability of $\beta_k < \beta_j$ between *class1* and *class2* for Pair (*k, j*) can be calculated by (2):

$$\begin{aligned} \Delta P_{kj} &= \left| P(\beta_k < \beta_j \mid class1) - P(\beta_k < \beta_j \mid class2) \right| \\ &= \left| P_{kj}(class1) - P_{kj}(class2) \right| \\ &\approx \left| \frac{a}{N1} - \frac{b}{N2} \right| \end{aligned} \tag{2}$$

For each Pair (*k, j*), another score was used to measure the average rank difference ($\Delta avgR$) between CpG site *k* and site *j* from *class1* to *class2*, which was calculated by (3):

$$\Delta avgR_{kj} = \left| \frac{\sum_{n=1}^{N1}(R_{n,k} - R_{n,j})}{N1} - \frac{\sum_{m=1}^{N2}(R_{m,k} - R_{m,j})}{N2} \right| \tag{3}$$

where *N1* and *N2* represent the number of profiles in *class1* and *class2*, respectively. $R_{n,k}$, $R_{n,k}$, $R_{m,i}$, $R_{m,j}$ represent the rank of site k (or j) in the *n*-th and *m*-th profile of *class1* and *class2* respectively.

C. Selection of R-pairs as markers

In the process of selecting marker R-pairs for each type of cancer, we first selected the *K* CpG sites with the highest appearance frequencies in all R-pairs. Then, for each of the *K* CpG sites, a CpG site was selected and paired to obtain a R-pair according to the following rules: for a CpG site *j* in *K* CpG sites, a site *i* was selected if the Pair (*i, j*) had the maximum $\Delta avgR$ score among all possible pairs composed of site *j*. If site *i* was in *K* CpG sites or had already been selected by other CpG sites in *K* CpG sites, then deleted Pair (*i, j*) from all possible pairs and selected the site according to the rule again.

D. Prediction rules

For a given methylation profile *X* , we can predict its class label by the following rules: first, for each selected marker R-pairs (*i, j*), if $P_{ij}(class1) > P_{ij}(classs2)$, then the class label of X can be assigned as (4)

$$X = \begin{cases} class1, & if\ x_i < x_j \\ class2, & otherwise \end{cases} \tag{4}$$

else if $P_{ij}(class1) \le P_{ij}(classs2)$, the class label of *X* can be assigned as (5)

$$X = \begin{cases} class2, & if\ x_i < x_j \\ class1, & otherwise \end{cases} \tag{5}$$

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

where x_i and x_j represent the methylation level of CpG site *i* and CpG site *j*, respectively. For *k* R-pairs, if more R-pairs vote *class1* than *class2*, then classified the profile *X* to *class1*, otherwise, classified this profile to *class2*. The specificity, sensitivity and accuracy defined as below were used to assess the prediction performance.

$$Specificity = \frac{TN}{TN + FP}$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN}$$

Where *TP*, *TN*, *FP* and *FN* indicate the number of true positive, true negative, false positive and false negative predictions, respectively.

III. RESULT

A. Selection R-pairs for five cancer types

For each training set for COAD, BRCA, KIRC, LUAD and STAD, we detected R-pairs with ΔP > 95% for STAD (only one R-pair was obtained with ΔP at 100%) and 100% for the other four cancer types respectively, resulting in 803395, 4674, 802, 10873 and 89 R-pairs, respectively. We found that, in each training dataset of five cancer types, the methylation levels of CpG sites involved in R-pairs were significantly higher than the methylation levels of CpG sites excluded in R-pairs (all p<=0.01, Wilcoxon ranksum test [11]; supplementary Fig.S1). The R-pairs involving the CpG sites with high appearance frequency pairs might be more suitable to be marker pairs. For example, as shown in Fig.1, the appearance frequencies of CpG sites involved in the R-pairs identified for COAD significantly correlated with the average differences of methylation levels between normal and tumor samples (Pearson coefficient: r=0.7574 and p<1.0×10-16). Thus, when we selected the top scoring pairs, which were composed of CpG sites with the highest appearance frequencies in all R-pairs (See method). For simplicity, the top 11 R-pairs (referred to as FR-pairs) involving 11 CpG sites with the highest appearance frequencies were selected as biomarkers for each cancer type. Then, the FR-pairs were tested on the testing set for each cancer type using majority vote rules. As shown in Fig.2a, the smallest percentage of correct classification was above 95%. Comparing with the top 11 R-pairs selected by the method of kTSP only according to ΔP and $\Delta avgR$ scores (Fig.2b), the FR-pairs had better prediction performance. Then, we compared the performance of the R-pairs to the performance of classifier constructed by traditional classify algorithms including k-nearest neighbors (KNN) and SVM. The classifier was constructed based on the top 100 most significantly methylated CpG sites between normal and tumor samples identified for each training dataset. Table SII showed that the performance of FR-pairs was better than the performance of KNN classifier and were comparable to the performance of SVM classifier on testing datasets (Table SI). Although the KNN and SVM classifier has good classification performance in distinguishing tumor and normal samples from the same data source (training sets and testing sets were all from TCGA) respectively, they may have worse transferability comparing to the R-pairs. The results indicated that FR-pairs could serve as effective signatures to distinguish cancer from normal samples.

B. Validation of FR-pairs

The FR-pairs identified for COAD could discriminate tumor samples from normal controls in dataset CRC44, which was from a different study, with a classification accuracy of 100%. Similarly, we also found that the FR-pairs identified for STAD and BRCA can discriminate the tumor samples in GAS75 and BRC248 from normal controls respectively with an accuracy more than 90% (Fig.3a). These results indicated that the R-pairs were insensitive to data source and FR-pairs might be transferable to different experimental datasets.

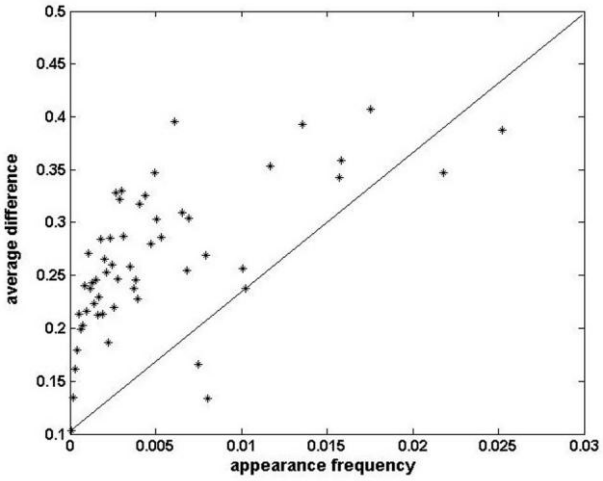

Fig. 1. Correlation between the appearance frequencies of CpG sites in R-pairs and differences of average beta-values between normal and cancer samples. X axis represents the frequencies of CpG sites involved in all detected R-pairs and Y axis represents the average differences of beta-values between normal and cancer samples for CpG sites having the same frequency in R-pairs.

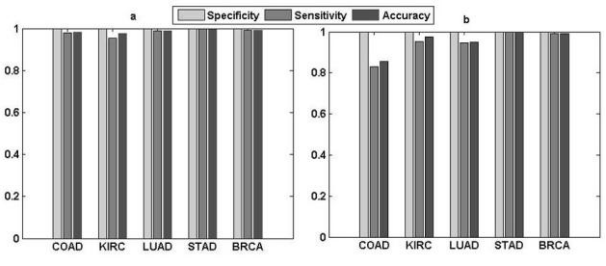

Fig. 2. Comparison of prediction performance between FR-pairs and R-pairs selected by k-TSP. a. The prediction performance of FR-pairs; b. The prediction performance of R-pairs selected by kTSP.

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

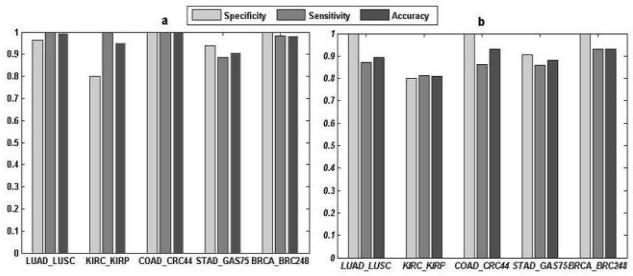

Fig. 3. Performance of R-pairs in independent datasets. a. The prediction performance of FR-pairs; b. The predication performance of R-pairs selected by kTSP.

Additionally, as shown in Fig.3b, the prediction performance of FR-pairs was better than R-pairs selected by kTSP in each cancer dataset.

The FR-pairs identified for a cancer type also performed well for different subtypes of the same cancer. For example, the FR-pairs selected for LUAD could also accurately predict the samples in Lung squamous cell carcinoma (LUSC), which is subtype of the non-small-cell lung carcinoma (Fig.3a), with the accuracy almost approaching 100%. For the validation dataset KIRP, which was composed of five normal samples and 16 kidney renal papillary cell carcinoma samples, only one normal sample was misclassified by FR-pairs obtained from KIRC. The results showed that the FR-pairs might grasp the common alterations of DNA methylation in different subtypes of the same cancer originating from different parts of the same tissue (LUAD and LUSC), or from different cell types (KIRC and KIRP).

The FR-pairs identified for a cancer type could discriminate samples of different degrees of malignancy for the same cancer type from normal controls. The BRC248 dataset includes 118 samples with the degrees of malignancy from stage 1 to stage 3 and 12 normal samples, which also includes 118 samples with four subtypes of breast cancer. The FR-pairs detected for BRCA showed a classification accuracy above 98% for all 248 samples (Fig.3a). The result validated that the FR-pairs could grasp common alterations of DNA methylation in different subtypes of the same cancer and also indicated that DNA methylation alterations might emerged in an early stage of cancer [12-14]. These results implied the FR-pairs might sever as biomarkers for early cancer diagnosis.

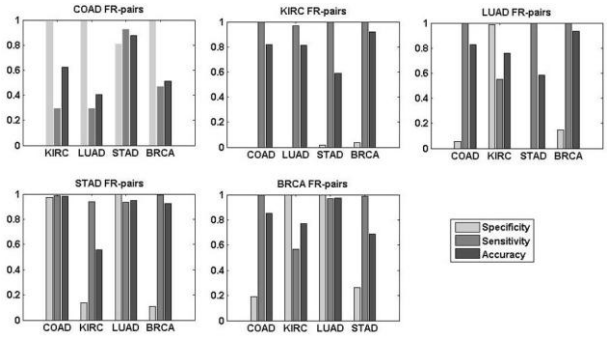

Fig. 4. Performance of FR-pairs in different types of cancer

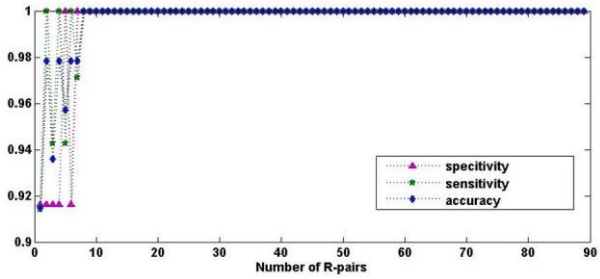

Fig. 5. The classification performance of multiple R-pairs detected from STAD by vote majority rule

Notably, the identified FR-pairs were specific to each cancer type. When using the FR-pairs identified for KIRC to classify the other four cancer type (STAD, COAD, LUAD, BRCA), the highest specificity is less than 20%. Similar results were observed for FR-pairs selected for each of the other four cancer types (Fig.4). These results indicated that DNA methylation alterations might have cancer specificity.

IV. Discussion

In this study, relative methylation level reverse pairs were detected from five cancer types respectively based on a simple rank based method. Especially, for each cancer type, the FR-pairs selected from the corresponding R-pairs performed well on discriminating normal and tumor samples. However, not all R-pairs were effective in classification. For instance, The R-pair (cg14242042, cg14012294), detected from the training set of COAD with only 24 samples, had poor classification performance in test dataset with specificity of 19%. Several factors could affect the performance. The small sample size in training set might lead to the identification of only a fraction of DNA methylation alterations in the general population. As sample size increases, such a R-pair might not be detected due to a low score of ΔP . The susceptibility of DNA methylation to environment, diet and aging might be another factor [15-16]. Additionally, other various biological and technical factors can also influence the relative order of two CpG sites. In particular, it has been reported that samples coming from different batches in datasets collected from TCGA might have batch effects[17]. Fortunately, using multiple R-pairs rather than a single pair may eliminate the influence caused by these factors. As shown in Fig.5, as the number of marker pairs increased the prediction performance increased. Specifically, when using more than 9 R-pairs detected from STAD, the accuracy approached 100%.

The FR-pairs could be biomarkers for cancer diagnosis as they could grasp the biological difference between normal and tumor samples. For example, the FR-pairs obtained from COAD involved 22 unique genes, which were all statistically significant (Wilcoxon rank sum test [11], with false discovery rate (FDR) [18] less than 0.05) differential methylated between normal and tumor samples. The methylation levels of CpG sites in column 1 of Table II were significantly up-regulated while in column 3 of Table II were significantly down-regulated when comparing tumor samples with normal samples.

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

TABLE II. FR-PAIRS IN COAD

Probe ID1	Gene symbol1	ProbeID2	Gene symbol2
cg15087147	LRRC4	cg26189983	TNFRSF1B
cg14242042	SOX5	cg14012294	BTF3
cg05714219	GALNT14	cg06627364	MGC4677
cg12874092	VIM	cg16120811	WEE1
cg04034767	GRASP	cg05445326	TM4SF19
cg07748540	PDGFD	cg25903497	TMBIM1
cg17872757	FLI1	cg06462291	NT5DC3
cg24446548	TWIST1	cg07903860	DCLRE1C
cg08190044	ATP8B2	cg08696192	TNS4
cg07570142	MOXD1	cg14547335	ATP2B2
cg05436658	PRKCB1	cg18289156	FLJ37357

Of the 11 CpG sites involved in FR-pairs (column 1 of Table II), six sites were annotated to genes (LRRC4, SOX5, VIM, FLI1, TVIST1 and PRKCB1) that have been reported to be hypermethylated in colon cancer and already considered as candidate biomarkers by other studies respectively [19-24]. As DNA methylation is sensitive to environmental influence, diet and aging and many other factors [15-16], different experiments may find different aberrant methylated CpG sites. For example, all CpG sites listed in Table II were found to be differentially methylated in COAD training dataset. However, three of them were not significantly differentially methylated in CRC44 dataset. When considering the relative methylation levels, we found that the ΔP of the three FR-pairs each involving one of the three CpG sites were 0.772, 0.818 and 0.909 in CRC44, indicating that these three CpG sites could be putative biomarkers in CRC44 dataset from the point of view of relative methylation levels. For another example, the specificity of KNN and SVM classifier constructed from BRCA on validation set BRC248, which were collected from GEO, were only 0.533 and 0.615 respectively. In contrast, the specificity of FR-pairs on BRC248 was 1, suggesting that the R-pairs may have higher transferable ability. This result indicated that it could be more stable to identify candidate biomarkers by considering relative methylation levels of two CpG sites than considering the individual CpG sites.

We have shown that the FR-pairs could capture the common alterations in different subtypes of the same cancer when compared with normal samples and had cancer specificity. In fact, different alterations in different subtypes of the same cancer types could also be detected by the rank-based methods. For instance, 841 R-pairs that could be detected between 127 LUAD and 133 LUSC tumor samples with cutoff of ΔP at 0.8, indicating the relative methylation levels of CpG sites involved in each of these R-pairs were reversal in at least 80% of LUSC tumor samples comparing to LUAD tumor samples. Similar result was also observed in KIRC and KIRP, where 28024 R-pairs were detected from 235 tumor samples composed of 219 KIRC and 16 KIRP samples, In the same way, we integrated the training set of five cancer types (Table SI) as a new training dataset to detect the common DNA methylation alteration patterns in different cancer types comparing with normal samples, for which 128 R-pairs were detected with ΔP > 0.8. Then, the new testing set, which involving 201 normal samples and 724 tumor samples, were classified by the top 11 of the 128 detected R-pairs with an accuracy of 97.5%. The result indicated that R-pairs can not only serve as biomarkers for cancer diagnosis but also detect common molecular alterations in cancer, which can shed light on the other follow-up studies such as drug target identification and mechanism of carcinogenesis.

ACKNOWLEDGMENT

This work was financially supported by the National Natural Science Foundation of China (grant numbers 91029717, 81071646, 81372213, 81201702 and 81201822).

REFERENCE

[1] Baylin S.B., S.A. Belinsky, and J.G. Herman, "Aberrant methylation of gene promoters in cancer---concepts, misconcepts, and promise," J Natl Cancer Inst, 2000. 92(18): p. 1460-1.

[2] Tost J., "DNA methylation: An introduction to the biology and the disease-associated changes of a promising biomarker," Mol Biotechnol, 2010. 44(1): p. 71-81.

[3] Dehan P., G. Kustermans, S. Guenin, J. Horion, J. Boniver, and P. Delvenne, "DNA methylation and cancer diagnosis: New methods and applications," Expert Rev Mol Diagn, 2009. 9(7): p. 651-7.

[4] Levenson V.V., "DNA methylation as a universal biomarker," Expert Rev Mol Diagn, 2010. 10(4): p. 481-8.

[5] Bhasin M., H. Zhang, E.L. Reinherz, and P.A. Reche, "Prediction of methylated cpgs in DNA sequences using a support vector machine," FEBS Lett, 2005. 579(20): p. 4302-8.

[6] Wang D., Q. Wang, F. Shan, B. Liu, and C. Lu, "Identification of the risk for liver fibrosis on chb patients using an artificial neural network based on routine and serum markers," BMC Infect Dis, 2010. 10: p. 251.

[7] Geman D., C. d'Avignon, D.Q. Naiman, and R.L. Winslow, "Classifying gene expression profiles from pairwise mrna comparisons," Stat Appl Genet Mol Biol, 2004. 3: p. Article19.

[8] Tan A.C., D.Q. Naiman, L. Xu, R.L. Winslow, and D. Geman, "Simple decision rules for classifying human cancers from gene expression profiles," Bioinformatics, 2005. 21(20): p. 3896-904.

[9] Barrett T., D.B. Troup, S.E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, "Ncbi geo: Archive for high-throughput functional genomic data," Nucleic Acids Res, 2009. 37(Database issue): p. D885-90.

[10] Bibikova M. and J.B. Fan, "Goldengate assay for DNA methylation profiling," Methods Mol Biol, 2009. 507: p. 149-63.

[11] Wilcoxon F., "Individual comparisons by ranking methods," Biometrics bulletin, 1945: p. 80-83.

[12] Brock M.V., C.M. Hooker, E. Ota-Machida, Y. Han, M. Guo, S. Ames, "DNA methylation markers and early recurrence in stage i lung cancer," N Engl J Med, 2008. 358(11): p. 1118-28.

[13] Guerrero-Preston R., A. Baez, A. Blanco, M. Berdasco, M. Fraga, and M. Esteller, "Global DNA methylation: A common early event in oral cancer cases with exposure to environmental carcinogens or viral agents," P R Health Sci J, 2009. 28(1): p. 24-9.

[14] Carmona F.J. and M. Esteller, "DNA methylation in early neoplasia," Cancer Biomark, 2010. 9(1-6): p. 101-11.

[15] Christensen B.C., E.A. Houseman, C.J. Marsit, S. Zheng, M.R. Wrensch, J.L. Wiemels, "Aging and environmental exposures alter tissue-specific DNA methylation dependent upon cpg island context," PLoS Genet, 2009. 5(8): p. e1000602.

[16] Johnson I.T. and N.J. Belshaw, "Environment, diet and cpg island methylation: Epigenetic signals in gastrointestinal neoplasia," Food Chem Toxicol, 2008. 46(4): p. 1346-59.

[17] Leek J.T., R.B. Scharpf, H.C. Bravo, D. Simcha, B. Langmead, W.E. Johnson, "Tackling the widespread and critical impact of batch effects in high-throughput data," Nat Rev Genet, 2010. 11(10): p. 733-9.

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

[18] Houseman E.A., W.P. Accomando, D.C. Koestler, B.C. Christensen, C.J. Marsit, H.H. Nelson, "DNA methylation arrays as surrogate measures of cell mixture distribution," BMC Bioinformatics, 2012. 13: p. 86.

[19] Koinuma K., R. Kaneda, M. Toyota, Y. Yamashita, S. Takada, Y.L. Choi, "Screening for genomic fragments that are methylated specifically in colorectal carcinoma with a methylated mlh1 promoter," Carcinogenesis, 2005. 26(12): p. 2078-85.

[20] Sinner D., J.J. Kordich, J.R. Spence, R. Opoka, S. Rankin, S.C. Lin, "Sox17 and sox4 differentially regulate beta-catenin/t-cell factor activity and proliferation of colon carcinoma cells," Mol Cell Biol, 2007. 27(22): p. 7802-15.

[21] Jung S., L. Yi, J. Kim, D. Jeong, T. Oh, C.H. Kim, "The role of vimentin as a methylation biomarker for early diagnosis of cervical cancer," Mol Cells, 2011. 31(5): p. 405-11.

[22] Kibriya M.G., M. Raza, F. Jasmine, S. Roy, R. Paul-Brutus, R. Rahaman, "A genome-wide DNA methylation study in colorectal carcinoma," BMC Med Genomics, 2011. 4: p. 50.

[23] Okada T., Y. Suehiro, K. Ueno, S. Mitomori, S. Kaneko, M. Nishioka, "Twist1 hypermethylation is observed frequently in colorectal tumors and its overexpression is associated with unfavorable outcomes in patients with colorectal cancer," Genes Chromosomes Cancer, 2010. 49(5): p. 452-62.

[24] Asting A.G., H. Caren, M. Andersson, C. Lonnroth, K. Lagerstedt, and K. Lundholm, "Cox-2 gene expression in colon cancer tissue related to regulating factors and promoter methylation status," BMC Cancer, 2011. 11: p. 238.

978-1-4799-7294-4/14/$31.00 ©2014 IEEE