Received: 6 November 2021 | Revised: 10 May 2022 | Accepted: 10 June 2022

DOI: 10.1002/cac2.12326

CANCER COMMUNICATIONS Open Access

LETTER TO THE EDITOR

DNA methylation signature predicts cancer response to demethylation agents from profiling diverse cancer cell lines

Dear Editor,

Abnormal DNA methylation, a process whereby tumor suppressors tend to be hypermethylated and silenced, is a hallmark of cancer cells [1]. Removing the methylation by demethylation agents such as azacitidine and decitabine is one of the strategies to treat cancers and has been successfully used to treat certain hematological and solid tumors [2]. However, there are no established DNA methylation markers or signatures that can accurately predict patients' response to treatment [2]; thus, the identification of reliable predictive biomarkers for effective therapy remains a critical need in clinical practice. Using genome-wide DNA methylation and response data to four demethylation agents (azacitidine, decitabine, RG108 and zebularine) in nearly 600 cancer cell lines, we systematically profiled the response patterns of the demethylation agents, conducted genome-wide association analysis of DNA methylation with the response for each drug, and identified key responsible pathways that could be associated with treatment response. Further, we applied machine learning techniques to develop a model to predict cancer's response to decitabine (Supplementary Materials and Methods, Supplementary Figure S1).

Our results showed that decitabine was the most potent drug among the 4 demethylation agents (smaller area under the curve [AUC] values represent greater drug potency), and hematopoietic/lymphatic cancer cells were the most responsive to all the drugs (decitabine and hematopoietic/lymphatic cancer cells had the smallest median or mean AUC; cancer cell line types were ordered by mean AUC ascendingly from left to right in Supplementary Figure S2A). These findings are consistent with clinical observations suggesting that cell lines could be representative of in vivo tumors. In the pair-wise comparisons between different drugs in all cell lines, decitabine demonstrated the smallest AUC response, zebularine was more effective than azacitidine, and no significant difference was observed between RG108 and zebularine or azacitidine (Supplementary Figure S2B).

In the genome-wide association analysis, we found a very large number of CpG sites significantly associated with decitabine and RG108 (102,183 and 90,826 at false discovery rate (FDR) < 0.05, respectively); however, there was almost none for azacitidine (0 CpG) and zebularine (1 CpG) (Figure 1A, Supplementary Figure S3). For the significant CpGs, about half of them were common between decitabine (46.7%) and RG108 (53.6%) (Supplementary Figure S4A) with identical positive or negative association directions (Supplementary Figure S4B). The vast majority of these CpGs (over 95%) were negatively correlated with drug response (Supplementary Figure S4C), i.e., hypermethylation of these CpGs was associated with better treatment response (smaller AUC). In addition, 79.2% and 88.6% of the CpG associated genes from decitabine and RG108, respectively, were also shared (Supplementary Figure S4D), which were enriched in 9 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Supplementary Figure S4E), such as calcium signaling, focal adhesion, mitogen-activated protein kinase (MAPK) signaling and Notch 1, as previously reported [3, 4].

We suspected that the hematopoietic and lymphoid cancer cell lines might have driven the decitabine response. Indeed, separate analysis for these cell lines alone showed that 75.8% of significantly associated CpGs were shared with the significant CpGs from all cell line analyses and that 8 out of 9 enriched pathways were common between the two (Supplementary Figure S5A-B). Stratified analysis for all other non-hematopoietic and lymphoid cancer cell lines demonstrated that DNA methylation was also significantly associated with decitabine response (1,4981 CpGs at FDR < 0.05 with 3,663 common with those found in

Abbreviations: AUC, Area Under the Curve; DNMT, DNA methyltransferase; FDR, False Discovery Rate; KNN, K Nearest Neighbors; RF, Random Forest; SVM, Support Vector Machine; TET, ten-eleven translocation.

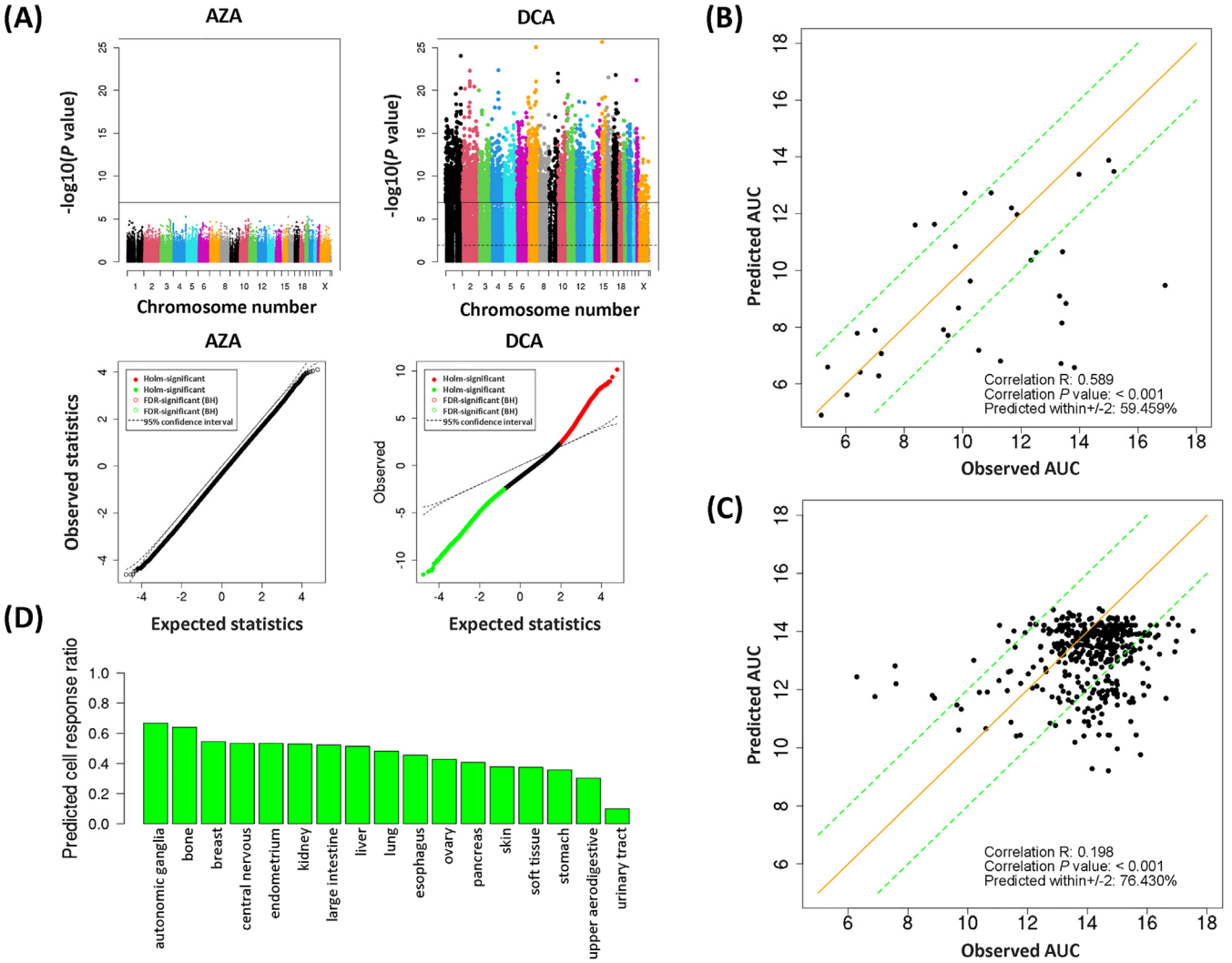

FIGURE 1 Association of CpG methylation with response and machine learning prediction. (A) CpG association with AZA and DCA. Top row: Manhattan plot for association *P*-value across chromosomes. No CpGs pass genome-wide significance for azacitidine; however, many were for decitabine. The horizontal solid line is the genome-wide significant cut-off, and the dashed line is adjusted *P*-value significance. Bottom row: Q-Q plot of CpG methylation and drug response. The diagonal line represents observed statistics (t statistics) is equal or similar to expected, i.e., there is no significant association between DNA methylation and drug response. More deviated from this line means more significantly correlated between the two. (B) Predicted vs. observed AUCs for the test dataset using KNN model in hematological and lymphoid cell lines (37 cell lines). The dashed green lines are +/-2 percent. (C) Predicted vs. observed AUCs for the independent test dataset consisting of all other non-hematological and lymphoid cell lines (489 cell lines). (D) Proportion of cancer cell lines (non-hematological and lymphoid) predicted to be responsive to decitabine in each cancer type. Abbreviations: AZA – azacitidine; DCA – decitabine; KNN – K Nearest Neighbors algorithm; AUC- Area Under the Curve

hematopoietic and lymphoid cancer cell analysis alone, Supplementary Figure S5C-D).

Using a 2:1 split ratio of training and testing in hematopoietic and lymphoid cell lines, we selected a subset of CpGs and developed machine learning models through three algorithms (K Nearest Neighbors [KNN], Random Forest [RF], and Support Vector Machine [SVM]). KNN was found to be the best performer in response prediction accuracy in the training set (data not shown). When using this model in the independent testing dataset of hematopoietic and lymphoid cell lines and the correlation between predicted and expected response scores was 0.589 ($P < 0.001$, Figure 1B), with about 60% of predicted scores within 2% of expected scores. This model was also applied to non-hematopoietic and lymphoid cell lines. Although the correlation between predicted and expected AUC scores was lower, it was also significantly correlated (correlation $R = 0.198$ and $P < 0.001$), with 76.43% of predicted scores within 2% of expected scores (Figure 1C). The proportion of cancer cell lines from different cancers predicted to be responsive ranged from 10% to 70% (Figure 1D), with cancers from the central nervous system and upper

aerodigestive tract having the highest proportion and were more likely to benefit from the treatment.

To determine the mechanisms of DNA methylation that could be associated with decitabine response, we examined the genome-wide methylation summary data in each cell line and correlated the summary data with cells' response to decitabine and mRNA expression of DNA methyltransferase (DNMT) and ten-eleven translocation (TET) methylcytosine dioxygenases. Although the hematopoietic and lymphoid cell lines had overall high DNA methylation (measured by mean or median, Supplementary Figure S6A), they were not the highest. They indeed had the highest variability (Supplementary Figure S6B), which may explain why only some patients responded to the treatment clinically. The mean DNA methylation was significantly correlated with decitabine response in both all cell lines (R = -0.193, Supplementary Figure S6C) and hematological/lymphoid cells alone (R = -0.297, Supplementary Figure S6D). Moreover, the mean DNA methylation was positively correlated with the expression of all three DNMTs (correlation coefficient R of 0.236, 0.227, and 0.244 for DNMT1, DNMT3A and DNMT3B, respectively, with P values all less than 0.001, Supplementary Figure S7A). Ironically, DNA methylation was also positively correlated with TET1 and TET3, although it was very weak (correlation coefficient R 0.109 and 0.172 with P values 0.002 and < 0.001, respectively, Supplementary Figure S7B).

Our investigation revealed several interesting findings. While many CpG sites were significantly associated with or predictive for the response to either decitabine or RG108, there were few for azacitidine or zebularine. The lack of association or predictability for azacitidine or zebularine suggests that baseline CpG methylation of cancer cells may not be a significant contributor to determining whether cancer cells respond to these drugs. However, it is possible that these drugs could act through other cells, such as reactivating immune cells' anti-tumor functions in the host [5–7]. The machine learning model for predicting decitabine response using DNA methylation data showed very good accuracy for hematopoietic and lymphoid cell lines. It also demonstrated high applicability to other cancer cells. The findings offer an expanded opportunity for many solid cancers (not used in current practice) with targeted demethylation therapy and suggest that decitabine acts on the common pathways between hematological and lymphoid cancers and solid cancers.

Despite the promising findings, certain limitations in this work should be mentioned. The data used were from consortiums, and the experimental conditions that might affect cell response or DNA methylation were not considered in the analysis. The results obtained from cell lines might not be similar in humans, where complex drug metabolism and tumor microenvironment could play significant roles in a tumor's response to treatment. Future studies with these issues in consideration, such as integrating experimental and multiple genomic data together and using in vivo model systems or tumor specimen methylation data, are needed to validate our results and possibly translate to practical clinical applications.

In summary, this large-scale study demonstrated that baseline DNA methylation of cancer cells predicts response to some (decitabine or RG108) but not all demethylation agents (azacitidine or zebularine). The findings showed that the machine learning approach could be a powerful tool to predict decitabine response by utilizing multiple CpGs. With further validation in tissue samples, machine learning may benefit patient selection with demethylation treatment.

DECLARATIONS

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All cell line data used in this study were publicly available, and no IRB approval or patient consent is necessary.

CONSENT FOR PUBLICATION

NA

ACKNOWLEDGMENTS

The authors would like to thank CTRP and GDSC for making the cell line genomic and response data available.

CONFLICT OF INTERESTS

The authors declare no competing interests.

FUNDING

This work was partly supported by the Mayo Clinic Center for Individualized Medicine.

AUTHOR CONTRIBUTIONS

ZFS and JPK conceived the research project. XWW and PV downloaded, prepared, and conducted part of the analysis. ZFS performed data analyses and drafted the manuscript. All of the authors read, edited, and approved the manuscript.

DATA AVAILABILITY STATEMENT

All data used in this study are publicly available from CTRP (https://portals.broadinstitute.org/ctrp.v2.1/?page=#ctd2BodyHome) and GDSC (https://www.ncbi.nlm.nih.gov/geo/; accession GSE68379) as described in the supplemental data.

Zhifu Sun
Xuewei Wang
Pete Vedell
Jean-Pierre Kocher

Division of Computational Biology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN 55905, USA

Correspondence
Zhifu Sun, MD, MS, Division of Computational Biology, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN 55906, USA.
Email: sun.zhifu@mayo.edu

REFERENCES

1. Pfeifer GP. Defining Driver DNA Methylation Changes in Human Cancer. Int J Mol Sci. 2018;19(4).
2. Dawson MA. The cancer epigenome: Concepts, challenges, and therapeutic opportunities. Science. 2017;355(6330):1147-52.
3. Ramakrishnan S, Hu Q, Krishnan N, Wang D, Smit E, Granger V, et al. Decitabine, a DNA-demethylating agent, promotes differentiation via NOTCH1 signaling and alters immune-related pathways in muscle-invasive bladder cancer. Cell Death Dis. 2017;8(12):3217.
4. Raynal NJ, Lee JT, Wang Y, Beaudry A, Madireddi P, Garriga J, et al. Targeting Calcium Signaling Induces Epigenetic Reactivation of Tumor Suppressor Genes in Cancer. Cancer Res. 2016;76(6):1494-505.
5. Daver N, Boddu P, Garcia-Manero G, Yadav SS, Sharma P, Allison J, et al. Hypomethylating agents in combination with immune checkpoint inhibitors in acute myeloid leukemia and myelodysplastic syndromes. Leukemia. 2018;32(5):1094-105.
6. Toor AA, Payne KK, Chung HM, Sabo RT, Hazlett AF, Kmieciak M, et al. Epigenetic induction of adaptive immune response in multiple myeloma: sequential azacitidine and lenalidomide generate cancer testis antigen-specific cellular immunity. Br J Haematol. 2012;158(6):700-11.
7. Lindblad KE, Goswami M, Hourigan CS, Oetjen KA. Immunological effects of hypomethylating agents. Expert Rev Hematol. 2017;10(8):745-52.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.