#238- Liquid biopsy to reveal colorectal cancer molecular subtype information from ctDNA epigenetics

Laura Herron[1], Patrick Cherry[1], Wen Zhou[1], April Choi[2], Vaia Florou[3], David Tsao[1]*

[1]BillionToOne, Inc., Menlo Park, Ca , [2]UC Irvine School of Medicine, Irvine, CA, [3] Huntsman Cancer Institute, Salt Lake City, UT

david@billiontoone.com*

INTRODUCTION

Circulating tumor DNA (ctDNA) is a valuable biomarker that supports therapy selection and on-treatment response monitoring through Northstar Select and Response, respectively. Northstar Response measures increases in tumor-specific DNA methylation, Tumor Methylation Score (TMS), via Quantitative Counting Template™ (QCT) technology to reveal tumor progression, enabling rapid treatment modification. During treatment selection, additional biological information provides prognostic value in predicting response, though is limited by tissue availability. The consensus molecular subtypes (CMS1-4)[1] of colorectal cancer (CRC) suggest potential immunotherapy response in CMS1 and enhanced benefit from irinotecan in CMS4. CMS1 accounts for 14% of CRC cases and is characterized by Micro-satellite instability (MSI-H), BRAF mutations, immune infiltration, and worse survival after relapse, while CMS4 accounts for 23% of cases and is characterized by worse overall survival, independent of relapse. In this analysis, we investigate whether ctDNA methylation patterns obtained from Northstar Response can reveal biological signal and subtype through a cohort-wide analysis. Successful identification of subtypes can help inform first line chemotherapy selection and inclusion criteria for clinical trials.

METHODS

We performed a retrospective analysis of 1710 blood samples from advanced cancer patients (Stage III, IV) that were submitted to BillionToOne's CLIA/CAP accredited clinical laboratory. Comprehensive DNA methylation analysis of >2000 differentially-methylated regions were performed using the Northstar Response assay. Unsupervised clustering analyses were then applied to ctDNA epigenetic features to identify molecular patterns. 321 colorectal samples had comprehensive cfDNA genomic profiling completed with Northstar Select and were used to generate hypothesis about the patterns observed in colorectal samples. All samples, besides non-cancer decoys, have a Tumor Methylation Score (TMS > 500).

RESULTS: Tumor Type Similarity

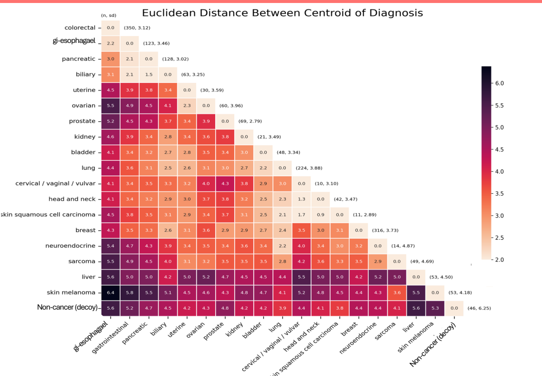


Fig. 2. Pairwise Euclidean distance between centroid of cancer types. The number of patients in each diagnosis group (n) and the within-group standard deviation (sd) are printed to the right of each row. The standard deviation of colorectal, pancreatic, and GI-esophageal groups is larger than their distance from each other. This suggests Response methylation data may be sufficient to determine the organ-system of disease.

RESULTS: GI Tumor Prediction

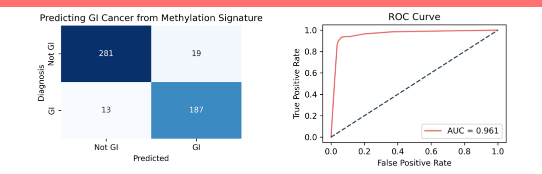


Fig. 3 Predicting if a patient's tumor is of GI origin using methylation data. A kNN model (n=11) is trained on the first 50 components of methylation data, labeled by whether they were diagnosed with a GI (colorectal, biliary, pancreatic, gastrointestinal) malignancy. The model has 93.7% specificity, 93.5% sensitivity, and 90.8% PPV for calling the tissue of origin of the tumor as GI.

RESULTS: Pan Cancer Methylation

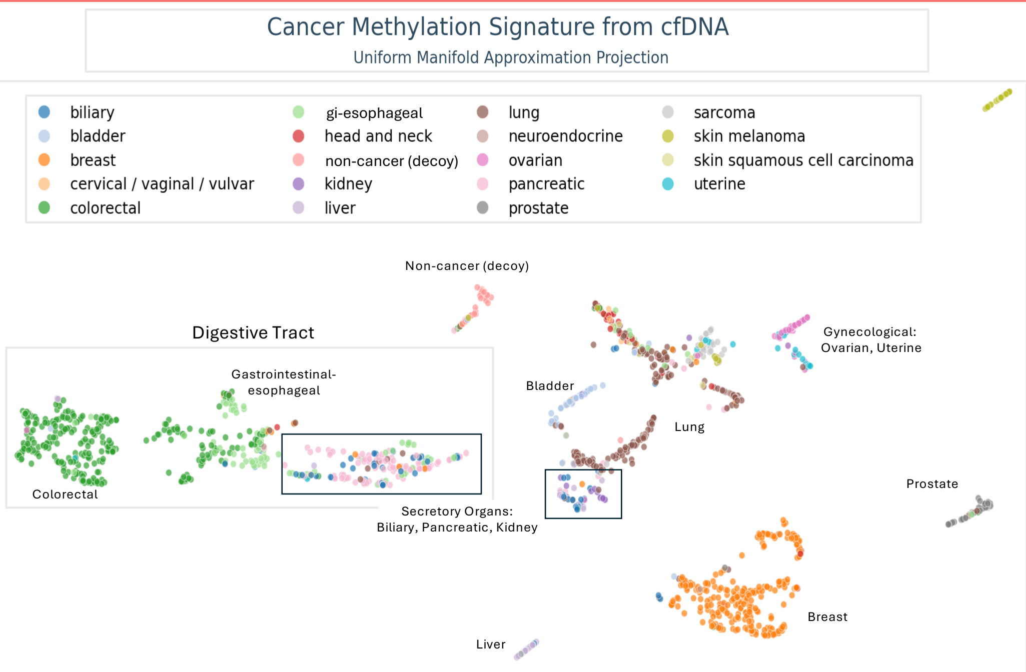


Fig. 1 Northstar Response Methylation Data Signatures Across Multiple Cancer Types. A. Uniform Manifold Approximation Projection (UMAP) on the first 50 principal components (PCs) of Response across ~2000 loci reveals biological signal. In addition to informing on tumor progression, this assay reveals tissue of origin groups that reflect organ systems and function. Within the digestive tract, colorectal and gastrointestinal-esophageal form unique groups, while the secretory organs (biliary, pancreatic, kidney) display overlapping signal. Methylation signal from non-cancer samples is used as a decoy control.

RESULTS: Colorectal Cancer Subtyping

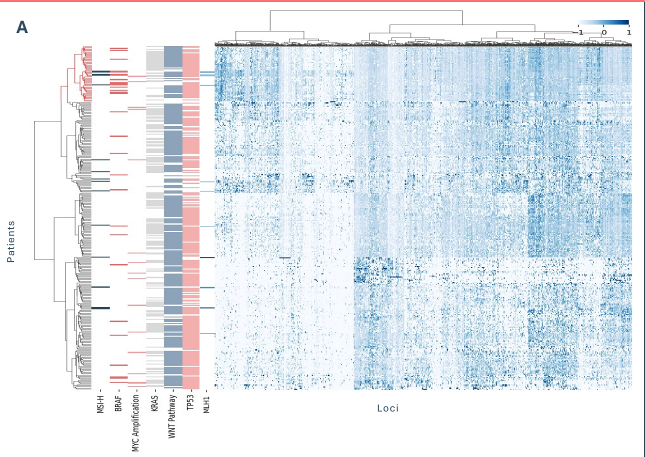
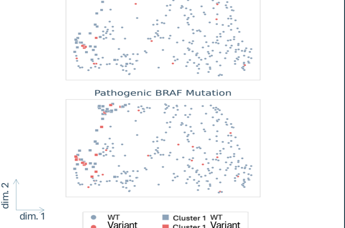


Fig. 4A Unsupervised clustering of 321 colorectal cancer methylation signatures. A. The heatmap of the methylation signatures (patients along *y* axis, loci along *x* axis) is annotated by the somatic mutational profile from Northstar Select. Key alterations[1] include MSI-H status, *BRAF*, *TP53*, and *KRAS* SNV, *MYC* amplification, WNT pathway (*APC*, *CDH1*, and *CTNNB1*) SNV, and *MLH1* methylation signal. Distinct banding patterns emerge with enrichment of MSI-H status and *BRAF* SNVs in the top cluster (red), which is characteristic of CMS1. *MLH1* promoter methylation is concordant with MSI-H status, as expected[3].

B

	Cluster 1	All Others	P-value
BRAF	15	14	2.15×10^{-6}
MSI-H	5	10	0.0711
MSI-H + BRAF	5	0	0.0001
Cluster Size	51	270	
	15.9%	84.1%	



B. Results of Fisher's exact test comparing the counts of *BRAF*/MSI mutational status of the highlighted cluster compared to the remaining data. *BRAF* mutational status was significantly different among the two groups (p = 2.15E-6). However, MS status is not statistically different (p = 0.071). The union of the two statuses (*BRAF* & MSI-H) was significantly different between the two groups, indicating overall enrichment of the CMS1 proxy type in the highlighted methylation profile cluster. **C.** The UMAP of the same methylation profiles shows a region of patients visually enriched with MSI-H/*BRAF*^{mut} on the bottom left. In hierarchical and graphical clustering, we observe all *BRAF*^{mut}/MSI-H patients co-locate to the same region, suggesting consistent methylation signatures.

DISCUSSION

The >2000 Northstar Response loci were designed for pan-cancer methylation tracking, but this analysis reveals that these loci also encode features of cancer biology that correspond to tumor of origin and molecular subtype. Unsupervised clustering analysis demonstrated distinct methylation signatures for the primary tumor, including origins from the GI-tract, melanoma, and liver. Methylation clustering of colorectal samples revealed 2 distinct groups. The first group (51/321, 15.9%), highlighted in red on Fig. 4A, is furthermore marked by a significant enrichment of BRAF^{mut} patients. Valenzuela *et al.* report that 14% of patients are the CMS1 subtype of whom 75% will be MSI-H and 40% will have *BRAF* mutations[2]. Additionally, CMS1 has CpG Island methylator phenotype (CIMP), and we observe increased methylation signal (darker) in the CMS1 band. Since tissue-based profiling was not performed for this real world study, RNA-based expression analyses to confirm the molecular subtype in these samples was unavailable. Nevertheless, the significant enrichment for BRAF alterations in methylation cluster 1 is consistent with CMS1. While MSI-H status is used as a current biomarker for immunotherapy response, we indicate a wider number of patients who may benefit. Cancer molecular subtyping has been shown to be a predictive biomarker therapy efficacy in several studies. A liquid biopsy for molecular subtyping could aid in therapy selection without the delays or impracticalities associated with RNA expression tissue profiling. The precision of Quantitative Counting Molecules™ allows us to distill biological meaning from methylation data and is inspiring future development.

ACKNOWLEDGMENTS

We would like to thank Oguzhan Atay for contributing scientific insight.

REFERENCES

1.Guinney, J., Dienstmann, R., Wang, X. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* **21**, 1350–1356 (2015).

2.Valenzuela G, Canepa J, Simonetti C, Solo de Zaldívar L, Marcelain K, González-Montero J. Consensus molecular subtypes of colorectal cancer in clinical practice: A translational approach. World J Clin Oncol. 2021 Nov 24;12(11):1000-1008.

3. Poynter JN, Siegmund KD, Weisenberger DJ, *et al.* Colon Cancer Family Registry Investigators. Molecular characterization of MSI-H colorectal cancer by MLHI promoter methylation, immunohistochemistry, and mismatch repair germline mutation screening. Cancer Epidemiol Biomarkers Prev. 2008 Nov;17(11):3208-15.