Poster 139

cfDNA Methylation Profiling Distinguishes Lineage-Specific Hematologic Malignancies

AACR Virtual Annual Meeting II
June 22-24, 2020

Qinwen Liu, PhD; Rita Shaknovich, MD, PhD; Xiaoji Chen, PhD; Zhao Dong, ScD, MS; M. Cyrus Maher, PhD, MS, MPH; Samuel Gross, PhD; Alexander Fields, PhD; Jan Schellenberger, PhD; Kathryn N. Kurtzman, MD; Eric T. Fung, MD, PhD; Anne-Renee Hartman, MD; Earl Hubbell, PhD; Arash Jamshidi, PhD; Alexander M. Aravanis, MD, PhD; Oliver Venn, DPhil

GRAIL, Inc., Menlo Park, CA

INTRODUCTION

- Hematologic malignancies, such as leukemia, lymphoma, and multiple myeloma, and their precursor conditions, such as monoclonal gammopathy of undetermined significance (MGUS) and monoclonal B-Cell lymphocytosis (MBL), are highly prevalent[1] as well as diverse in biology.[2,3]
- Hematologic malignancies arise as a result of deregulated differentiation in hematologic cell lineages.[4] A blood test that accurately and simultaneously detects and distinguishes the various types of hematologic conditions could provide a convenient approach to guide diagnostic workup and treatment selection.
- Recent results from the second sub-study of the Circulating Cell-free Genome Atlas study (NCT02889978) demonstrate that a targeted methylation- and cfDNA-based multi-cancer early detection (MCED) test detects >50 cancers with high specificity (>99%) and accurately predicts the tissue of origin (TOO) in >90% of eligible cases.[4]
- Here, we applied the same targeted methylation-based technology to refine cancer detection and TOO prediction accuracy for five major hematologic malignancies, and to identify methylation signatures specific to the developmental lineages of these conditions.

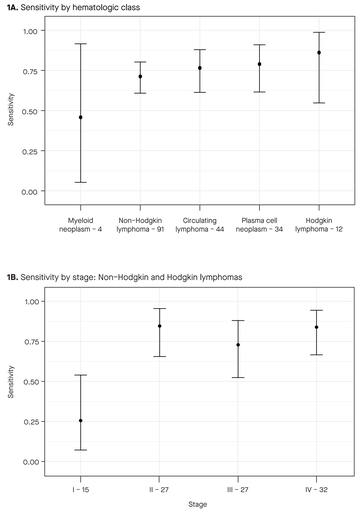
METHODS

- cfDNA samples were from the second pre-specified sub-study of CCGA, which was designed for targeted methylation assay validation. During the study period, only training set samples were available and used for analysis. For classification model training, we also included tumor tissue samples from an in-house tissue biopsy reference database.
  - Samples used to train the custom classification model for hematologic malignancies were from participants enrolled with a hematologic cancer diagnosis (cancer cases; n=301) and participants enrolled without a cancer diagnosis (non-cancer controls; n=2,687). 154 blood cell samples or tissue FFPE samples of hematologic malignancies were also included.
  - Classification performance was evaluated on cfDNA samples not previously used for targeted methylation panel design.[5] In total, cfDNA samples from 185 participants with hematologic cancers and 1,998 non-cancer controls confirmed without cancer diagnosis at the one year follow-up were included for performance evaluation.
- Five hematologic classes were used as training labels for classifier training: myeloid neoplasm, circulating lymphoma, Hodgkin lymphoma, Non-Hodgkin lymphoma, and plasma cell neoplasm.
- A cross-validated mutual information-based algorithm was used to identify features that discriminated between the five hematologic classes and the control class.
- A multinomial classifier was then trained to detect the presence or absence of cancer at the target specificity of 99.5% and predict TOO among the five major hematologic cancers and non-cancers using six-fold cross-validation.
- A low dimensional representation of the methylation features active for the final classifier was generated using the UMAP method,[6] which preserves the topology of high dimensional data.
- Scatter plots were used to visualize the relationship between the spatial distribution of feature embeddings and the model's class probabilities.

RESULTS

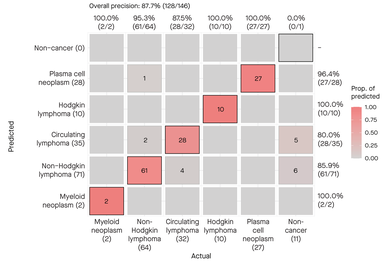
- The hematologic-specific classifier achieved an overall specificity of >99.4% [95% CI: 99.0-99.7%] in controls, and a sensitivity of 74.3% [67.4-80.5%] for hematologic cancers.
- The sensitivities by hematologic classes were 45.8% [5.3-91.6%] for myeloid neoplasms, 76.5% [61.3-88.0%] for circulating lymphomas, 86.1% [54.7-98.7%] for Hodgkin lymphomas, 71.3% [60.8-80.3%] for other Non-Hodgkin lymphomas, and 78.9% [61.6-91.0%] for plasma cell neoplasms (**Figure 1A**). For lymphomas (Hodgkin lymphomas and Non-Hodgkin lymphomas), the sensitivities by stages were 25.6% [7.2-54.0%] for stage I, 84.6% [65.5-95.5%] for stage II, 72.8% [52.4-88.0%] for stage III, and 83.9% [66.6-94.4%] for stage IV (**Figure 1B**).

Figure 1. Hematologic Classifier Sensitivity at the 99.5% Target Specificity Level. The number in the class label indicates the number of samples. The error bars show the 95% confidence intervals. (A) Sensitivity by hematologic class. (B) Sensitivity by stage for Hodgkin and Non-Hodgkin lymphomas.



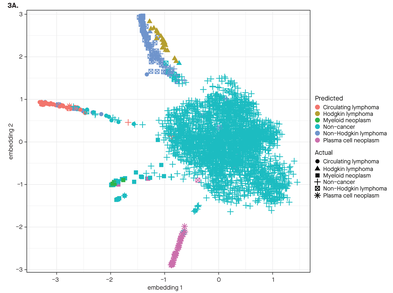
- TOO prediction was assessed on cancer cases that were correctly detected by the classification model. The hematologic-specific classifier achieved an overall TOO prediction accuracy of 87.7%, with Hodgkin lymphoma and myeloid neoplasm showing the highest prediction accuracy (100%) followed by plasma cell neoplasm (96.4%), Non-Hodgkin lymphoma (85.9%), and circulating lymphoma (80%, **Figure 2**).

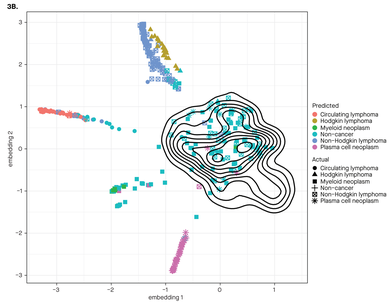
Figure 2. Tissue of Origin Prediction for Cases Predicted As Hematologic Cancers. Numbers in each box represent the total number of predicted TOO. Color corresponds to the proportion of predicted TOO, as indicated to the right of the plot. Percent correct predictions from the total predictions for each cancer type are indicated to the right of the plot.



- Of the 11 non-cancer controls (0.55% of non-cancer controls) who were classified as having hematologic cancers, five were predicted as circulating lymphoma, and six were predicted as other Non-Hodgkin lymphoma (<1% false positive rate), most showing confident TOO signal localizing to the predicted heme class (≥50% of total probability mass).
- Proximity in low-dimensional UMAP space (embedding dimensions) of input features showed that the majority of hematologic malignancies separated into five major clusters reflecting developmental lineages and disease ontogeny (**Figure 3A**). The vast majority of non-cancer controls were clustered separate from the hematologic cancers (**Figure 3B**).

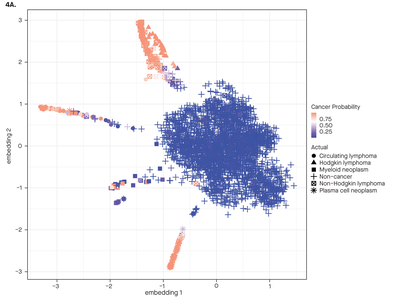
Figure 3. Hematologic Malignancies Are Clustered Into 5 Major Groups. Classification features were visualized using the UMAP method. (A) Hematologic malignancy cancer samples visualized by predicted versus actual TOO label. The shape of the symbol represents the actual hematologic class labels. The color of the symbol represents the predicted class label. (B) The same plot except that non-cancer controls are shown using contour density to reveal the rest of hematologic cancer samples.

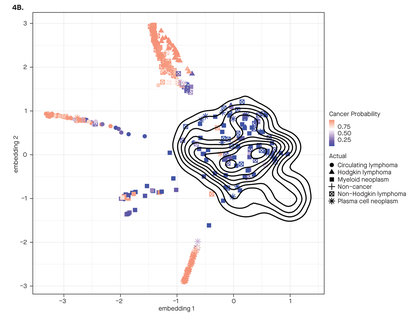


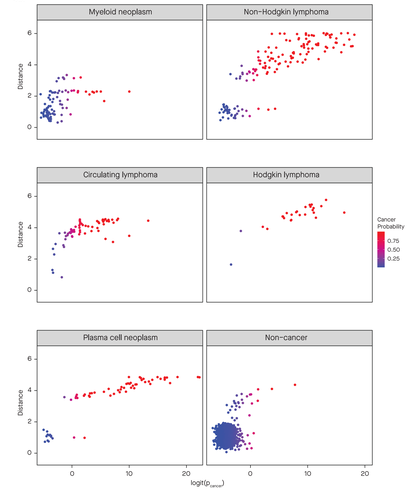


- The position of samples within each hematologic cluster correlated with the classification score. For individual cases with significant cancer signal (classification score >0.5), which are shown as light blue to red symbols in **Figure 4A and 4B**, there was strong positive correlation between their UMAP embedding localization and classification score (**Figure 4C**).

Figure 4. The Correlation Between Classification Scores and the Spatial Localizations on UMAP Embedding Space. (A) Hematologic cancer samples visualized by predicted versus actual TOO label. The color of the symbol represents the cancer probability of the sample. The shape of the symbol represents the class label of the sample. (B) The same plot except that non-cancer controls are shown using contour density to reveal the rest of hematologic cancer samples. (C) The correlation between the localization and the classification score by participant cancer status. The x-axis is the logit transformed probability of a sample being cancer. The y-axis is the Euclidean distance of a sample to the centroid of the non-cancer population in the UMAP embedding space. The non-cancer centroid is computed by averaging each embedding dimension of all non-cancer samples.







CONCLUSIONS

- Methylation features of cfDNA in patients with hematologic malignancies delineated five major clusters that reflected hematologic lineages.
- Lineage-specific signals followed a gradient suggestive of variation in disease-related methylation or tumor DNA shedding.
- The custom classifier for hematologic malignancies offers a convenient way to simultaneously detect and distinguish five major hematologic malignancies, which could help facilitate clinical diagnosis and treatment selection.
- Most cfDNA arises from circulating blood cells; therefore, characterizing methylation changes in conditions that are precursors of hematologic cancer may be important for improving performance of the multi-cancer early detection test. The findings from this study will guide further efforts toward removing interfering biological signals from cfDNA-based cancer detection assays and achieving even more sensitive detection of multiple cancers.

References

1. Li J, et al. Estimating the prevalence of hematological malignancies and precursor conditions using data from Haematological Malignancy Research Network (HMRN). *Cancer Causes Control*. 2016;27:1019-1026.
2. Arber DA, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127:2391-2405.
3. Swerdlow SH, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127:2375-2390.
4. Bowman RL, et al. Clonal Hematopoiesis and Evolution to Hematopoietic Malignancies. *Cell Stem Cell*. 2018;22:157-170.
5. Liu MC, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Ann Oncol*. 2020;31:745-759. DOI:https://doi.org/10.1016/j.annonc.2020.02.011.
6. McInnes L, et al. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv:1802.03426 [stat.CS]*. 2018.

Disclosures

All authors are employees and stockholders of GRAIL, Inc. AJ also holds equity in Illumina, Inc. AMA is an advisor to and holds equity in Foresite Labs, LLC.

Acknowledgments

Assistance with creative layout of presentation was provided by Robert Shaddock of ProEd Communications, Inc. (Beachwood, OH) and was funded by GRAIL, Inc.

Corresponding author

Qinwen Liu, qliu@grailbio.com