9064W

Using Cell Type Deconvolution to Add Immune, Metastatic, and Health Context to Methylation-based Early Cancer Detection

Harbinger Health

Kristy Schlueter-Kuck[1], Shiva Farashahi[1], Yifan Wu[1], Boyang Zhang[1], Miguel Williams[1], Tony Shuber[1], Kieran I Chacko[1], Jocelyn Charlton[1*]
[1]Harbinger Health, Cambridge, MA *Corresponding author: jcharlton@harbinger-health.com

BACKGROUND

Methylation-based liquid biopsy assays using cell free DNA (cfDNA) show promise for multi-cancer early detection (MCED). While cfDNA is a mixed population from multiple cell types, we have previously demonstrated high sensitivity in detecting rare tumor-derived signal and accurately assigning the tissue of origin[1]. However, the majority of cfDNA originates from blood cells and other normal tissues, which have the potential to offer insight into an individual's health, immune status, or comorbidities as well as improve contextualization of cancer-detection models and test results. Here, we present a cell type deconvolution (CTD) method and apply it to cfDNA and white blood cell (WBC) samples. Importantly, our approach utilizes our existing MCED cancer-detection assay (HHx) that interrogates 18.6 Mb of the genome[2] generating novel information that can add value to reported findings.

METHODS

We applied a previously published CTD algorithm[3] to our HHx assay. Our input reference data included public whole genome bisulfite sequencing from sorted cells from (1) the IHEC Blueprint Atlas[4] (28 blood cell types, n=94), and (2) the Atlas of Normal Cell Types[5] (82 blood and healthy tissue cell types, n=207). We supplemented these data with 134 cancer biopsy samples from 8 indications processed through the HHx assay to enable detection of cancer signal (biliary, esophagus, head & neck, gallbladder, liver, lung, pancreas, stomach). As biopsy samples are heterogenous, we computationally enriched the tumor purity using a novel read filtering approach[1]. To select informative features, we subset reference data to CpGs covered by our assay (N = 1,064,387) and used a two-sample t-test for each cell type to identify CpGs with significantly different mean methylation values from those of other cell types. We then calculated cell type proportions for 21 cell types across 2,541 cfDNA samples (NCT05435066, 619 cancer, same indications as above and 1,922 non-cancer) of which 1,363 also had matched WBCs (330 cancer, 1,033 non-cancer) processed by our HHx assay. Absolute concentrations were calculated by adjusting for cfDNA and plasma yield and then multiplying by a scaling factor of 3.22 pg per haploid genome. Results from a complete blood count with differential were available for 161 cancer subjects with matched WBC samples.

RESULTS

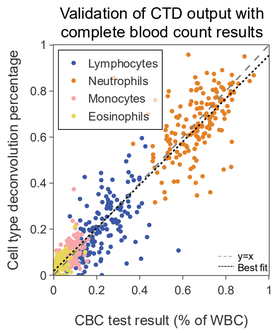

Figure 1. Correlation between complete blood count results and WBC CTD proportions.

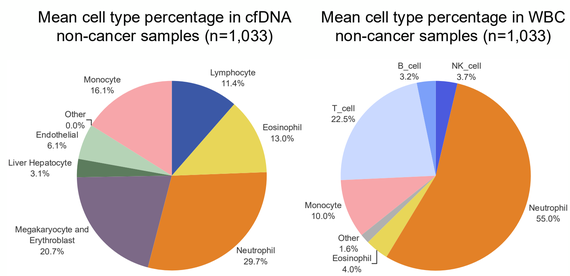

Figure 2. Pie charts showing the mean cell type percentages in cfDNA (left) and WBC (right) samples for 1,033 non-cancer subjects.

Cancer signal detection

To assess our ability to detect cancer signals, we compared the CTD estimates for the 619 cfDNA cancer samples to each sample's estimated tumor fraction. Here, we observed indication-specific signal, for example cases with non-small cell lung cancer or head and neck cancer showed increased proportions of those cancer cells, while the majority of other cancer types and non-cancer samples remained undetectable (**Figure 4**). This suggests that using cancer biopsy samples within the reference atlas may be an effective strategy for predicting cancer type. For 13 subjects, we detected liver cancer signal despite the subjects having a non-liver primary cancer type (**Figure 4**). Interestingly, 5/13 were stage IV cancers with liver metastases, hence the detected signal may be linked to shedding of local tissue at the metastatic site (**Figure 4**). However, we do not observe this effect in samples with lung metastases. As our model does not currently quantify proportions of DNA from healthy lung tissue, further iteration may improve metastasis identification.

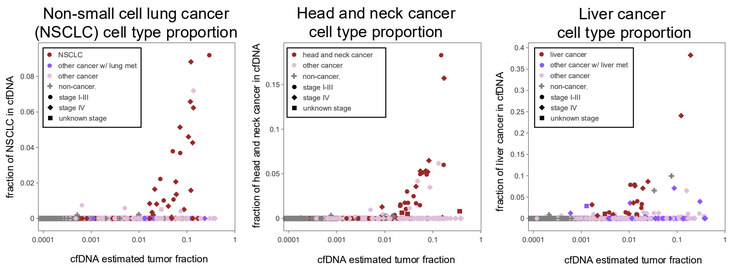

Figure 4. CTD-based detection of cancer cfDNA. Comparison of estimated tumor fraction and proportion of cfDNA attributed to non-small cell lung cancer, head and neck and liver cancer cells.

Cell type proportions in WBCs and cfDNA

Our WBC estimates correlated with complete blood count data (Pearson's r=0.95, **Figure 1**) and both cfDNA and WBC cell type proportions were consistent with expected biological profiles, including the absence in WBCs of megakaryocyte and erythroblast cells (cell types compartmentalized to the bone marrow) as well as liver hepatocytes and endothelial cells (**Figure 2**). Of interest, for cell types present in both cfDNA and WBCs, estimated proportions did not correlate between matched samples (r = 0.16 to 0.43). The lack of correlation suggests that the amount of cfDNA released from WBCs (or the cell turnover rate) may not be consistent across individuals. We observed wide variability in cell type proportions across individuals for both cfDNA and WBCs, with neutrophil and megakaryocyte/erythroblast cell types showing the greatest variability in non-cancer samples (**Figure 3**). Furthermore, we detected the presence of cancer cell types specifically within the cancer cfDNA cohort (**Figure 3**).

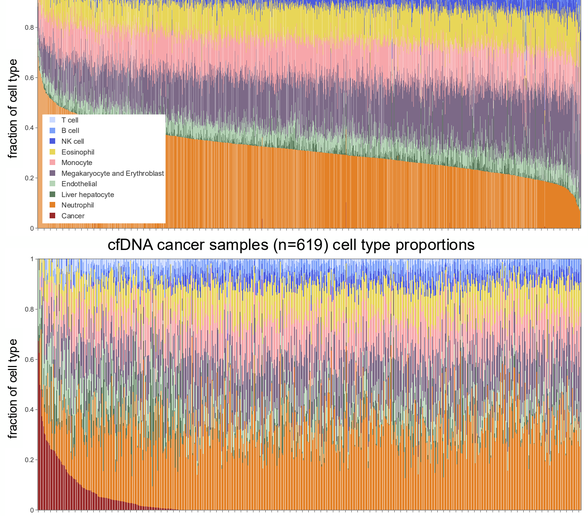

Figure 3. Variation in cell type proportions across non-cancer and cancer cohorts. Each column is a sample; non-cancer samples are sorted by neutrophil cell type percentage and cancer samples are sorted by the summation of CTD cancer cell fractions.

RESULTS

Immune modulation

For 11.3% cancer samples (70/619) and 0.1% non-cancer samples (2/1,922), we observed a complete dropout of NK cell-derived cfDNA (<0.05 genomic equivalents per mL plasma) (**Figure 5**) despite non-zero counts in WBCs.

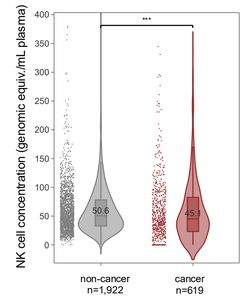

Figure 5: cfDNA NK cell concentration. By Wilcoxon rank sum test, ***: p<0.001.

To assess the immune state, we computed the neutrophil-to-lymphocyte ratio (NLR) in WBCs and observed a statistically significant increase from subjects with no reported ongoing medical conditions to subjects with ongoing immune, autoimmune, inflammatory, or blood disorders, to those with cancer (**Figure 6**).

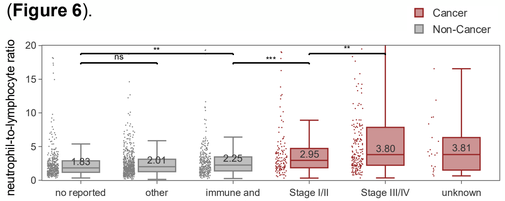

Figure 6: WBC neutrophil-to-lymphocyte ratio across a spectrum of disease states. Outliers with NLR>20 are not visualized for clarity. By Wilcoxon rank sum test, ns: not significant, **: p<0.01, ***: p<0.001.

CONCLUSIONS

- We present a methylation-based CTD approach and apply it to a large cohort of non-cancer and cancer subjects.
- cfDNA and WBC composition are highly variable among individuals and may reflect an individual's immune health.
- This method shows potential to aid in tissue of origin determination including informing on metastatic disease.
- Overall, our CTD algorithm provides additional insights that complement a MCED assay.

REFERENCES

1. *Farashahi, S., Wu, Y. et al. Denoising Models Enhance Detection of Tumor-Derived cfDNA fragments and Cancer Tissue signal in Liquid Biopsy. Proceedings of the AACR Special Conference in Cancer Research: Artificial Intelligence and Machine Learning; (2025).*
2. *Gregg, J., Michor, F. et al. Novel blood-based assay for detection of early stage multi-cancer. J Clin Oncol 41, e15035-e15035 (2023).*
3. *Houseman, E.A., Kile, M.L. et al. Reference-free deconvolution of DNA methylation data and mediation by cell composition effects. BMC Bioinformatics 17, 259 (2016).*
4. *Adams, D., Altucci, L. et al. BLUEPRINT to decode the epigenetic signature written in blood. Nat Biotechnol 30, 224–226 (2012).*
5. *Loyfer N., Magenheim J. et al. A DNA methylation atlas of normal human cell types. Nature (2023); 613:355–64.*

DISCLOSURES

This study was sponsored by Harbinger Health. Cambridge, MA

ACKNOWLEDGEMENTS

Editorial support was provided by Sean Husick (Harbinger Health, Cambridge, MA)
We gratefully acknowledge all participants for their contributions, without whom this research would not have been possible.

www.harbinger-health.com