LB254

# Sharpening the Signal: Enhancing Liquid Biopsy Specificity Through Intra-Individual Methylation Analysis

Harbinger Health

*Miguel Williams, Sarah Falotico, Caitlin Gilley, Mayur Gurnani, Emily Neaga, Kristy Schlueter-Kuck, Shiva Farashahi, Jackson A. Killian, Jocelyn Charlton, Kieran Chacko, Tony Shuber*

## BACKGROUND

Early cancer detection has consistently been associated with better patient outcomes. However, current detection methods are underutilized and often unscalable for population wide implementation. Liquid biopsy multi-cancer early detection (MCED) has emerged as a potential solution to accessibly detect cancers, but it is technically challenging to identify scant disease signals while differentiating from background biological signal. Cell free DNA (cfDNA) methylation has been identified as an excellent biomarker for use in liquid biopsy MCED tests. While circulating tumor DNA (ctDNA) often comprises less than 0.1% of cfDNA, there are often more methylation disease loci than mutation disease loci within a cancer genome. The higher abundance of methylated loci increases the odds that ctDNA is observed in a liquid biopsy MCED due to stochastic sampling loss. Harbinger Health's proprietary methylation biomarkers have been consistently observed across cancer indications and stages, suggesting that they are present at the start of cancer initiation. Physiological conditions such as aging can also change methylation patterns; these sporadic methylation fluctuations can be construed as ctDNA, impeding accurate disease identification.

Approximately 60% of cfDNA originates from white blood cells (WBCs).[1] We hypothesized that a paired intra-individual analysis (IIA)—comparing plasma-derived cfDNA to matched WBC-derived genomic DNA (gDNA)—could help differentiate ctDNA signal from background somatic noise (see Figure 1). To evaluate this hypothesis, blood draws were processed on subjects from the CORE-HH clinical study *(NCT05435066)*.

## METHODS

Subject blood draws from 1531 subjects were fractionated and cfDNA and gDNA were independently extracted from plasma and WBCs, respectively. WBC gDNA samples were sheared prior to downstream processing. Both DNA types were independently bisulfite treated, converted into indexed libraries, enriched for cancer specific regions of interest and sequenced to ≥100X unique median target coverage. Of the 1531 subjects, 1280 were evaluable by machine learning (ML) classifier, comprised of 943 patients with no reported cancer diagnosis, and 337 with a confirmed cancer diagnosis from 9 cancer types (Lung, Head & Neck, Pancreas & Gallbladder, Stomach, Esophagus & GEJ, Liver & Biliary).

IIA data was analyzed by four methods. First, a cfDNA-only ML cancer classifier was built (MLX) (see Figure 2A). Second, a patient-specific intra-individual classifier (IIX) was built by subtracting the fraction of tumor content (TC) signal observed from WBC from the fraction of TC signal detected in cfDNA (see Figure 2B). A 10-fold cross validation (CV) approach was used to establish specificity thresholds for both MLX and IIX. Subjects were called as cancer when determined positive by both IIX and MLX (See Figure 3). Third, a read-level IIA-filtering analysis identified and removed reads with shared phased methylation patterns in both patient samples, with the goal of enriching the cfDNA for ctDNA (see Figure 2C). Fourth, cell type deconvolution was performed on WBC data to identify the composition of each WBC subtype.

## RESULTS

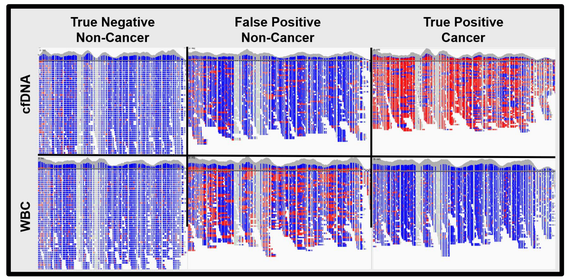


**Figure 1: Genome visualization of exemplary IIA comparisons.** Visualized NGS read stacks. Unmethylated CpGs are colored blue and methylated CpGs are colored red. True negatives as detected by cfDNA are expected to have low methylation across both cfDNA and WBC. False positives are expected to have elevated background WBC methylation levels which drive elevated cfDNA methylation. True positives are expected to have high cfDNA methylation levels compared to WBC due to methylated ctDNA.

| Tiers | MLX (Target Spec.) | IIA (Target Spec.) | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|---|---|
| 1-Tier | cfDNA MLX (98.5%) | - | 68.249% | 98.727% | 30.193% | 99.741% |
| | - | Paired IIX (98.5%) | 66.469% | 98.515% | 26.528% | 99.726% |
| 2-Tier | cfDNA MLX (98.5%) | Paired IIX (98.5%) | 56.973% | 99.470% | 46.425% | 99.652% |
| | cfDNA MLX (99.5%) | Paired IIX (99.5%) | 55.192% | 99.893% | 80.759% | 99.639% |

**Table 1: MLX and IIX Performance Results.** Performance results of the N1280 subject cohort each independent assay (1-tier) and both assays combined (2-tier). The combination of both assays raised the achieved specificities above those of the target specificities.

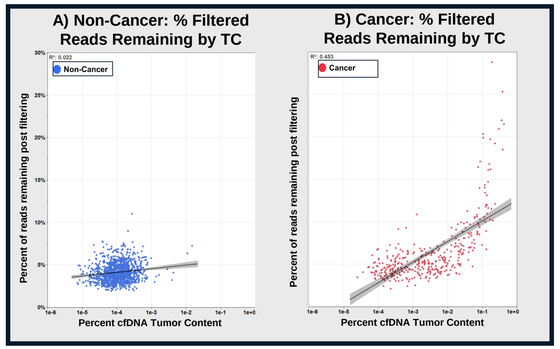


**Figure 5: Post IIA read level filtering anomalous read fractions. A)** Non-Cancer patients had few cfDNA reads post filtering. **B)** Cancer patients cfDNA reads post filtering correlated with cfDNA TC.

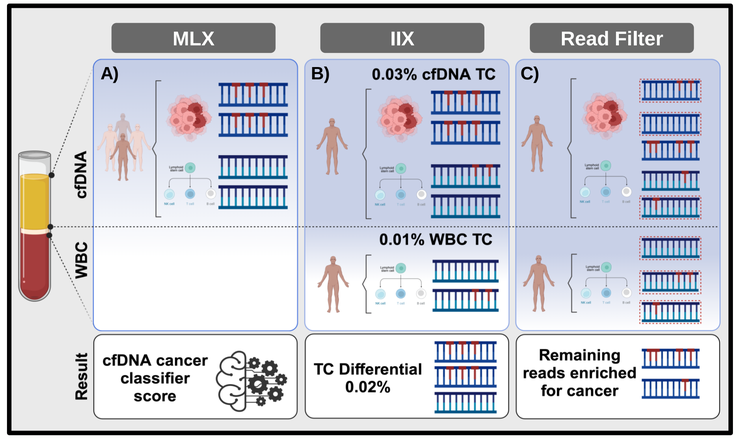


**Figure 2: IIA workflow diagram. A)** MLX defines cancer status based on cfDNA. **B)** IIX is used to differentiate cfDNA tumor fraction signal from WBC derived methylation which is conflated as tumor fraction. **C)** Read filtering compares cfDNA and WBC data and removes reads which have similar methylation patterns in both datasets.

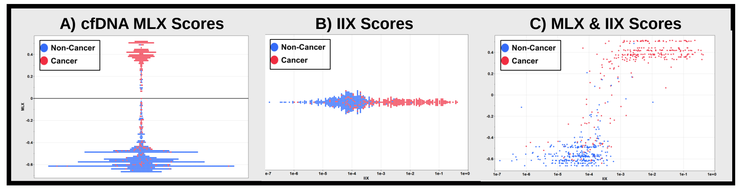


**Figure 3: MLX and IIX N1280 Data.** A two-tier procedure was utilized to evaluate cancer classification. Subjects who were defined as assay positive by both. **A)** the cfDNA derived MLX, and **B)** the cfDNA WBC comparative IIX, were called as cancer positive **C)**.

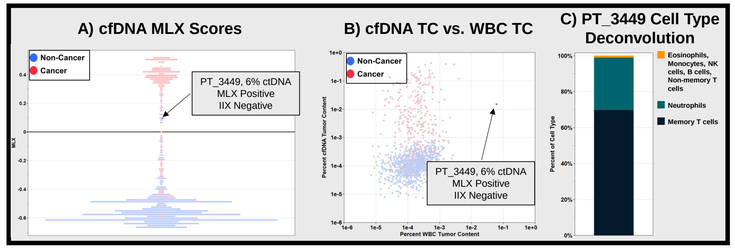


**Figure 4: IIX score and cell type deconvolution of PT_3449. A)** Subject PT_3449 had no suspicion of cancer, but was called positively by MLX. **B)** PT_3449 had high cfDNA and WBC TC. **C)** Cell type WBC deconvolution identified that PT_3449 had a high fraction of Memory T cells, suggestive of a possible immune response.

## DISCUSSION

IIA was performed on increasingly large cohorts, comprised of 36, 335 and lastly 1531 subjects. In all instances, performing IIX resulted in lower average scores for non-cancer subjects than for cancer subjects. When MLX and IIX were used independently on the 1280 samples at a 98.5% target specificity, sensitivities of 68.2% and 66.5% were achieved, respectively (see Table 1 and Figure 3). When the MLX and IIX were combined (i.e. only IIX-pass samples were evaluated by MLX), they achieved a 63.7% sensitivity at an achieved specificity of 99.5% which generates a PPV of 54.8%. While variability inherently increases at extreme values, when IIX and MLX were both set to target specificities of 99.5% the achieved sensitivity and specificity was 55.1% and 99.89% with a PPV of 80.7%.

One representative subject, PT_3449, with no reported cancer, was positive by MLX but then called negative by IIX (see figure 4a). The data from this subject was an extreme outlier, it had both the highest percentage of ctDNA (1.49%) and WBC TC (6.14%) of any non-cancer sample (see Figure 4B). Cell type deconvolution indicated that approximately 70% of the WBCs in this subject were comprised of memory T cells, one of the largest fractions observed in any of the subjects in this cohort (see Figure 4C). This fraction of memory T cells is well above the expected range which is between ~10-25%.[2] The high portion of memory T cells suggests a possible immune reaction at the time of blood draw.

After IIA read filtering, whereby reads with similar methylation patterns (ie. >99% probability that the methylation patterns are the same across both cfDNA and WBC data) were removed. Non-cancer subjects were found to have fewer remaining reads (i.e. presumptively ctDNA enriched), than cancer subjects (see Figures 5A and 5B). The fraction of remaining reads post filtering was weakly correlated with cfDNA TC in cancer sujects, but not correlated in non-cancers subjects. This relationship between filtered reads and TC further suggests that WBC can effectively identify background signal.

## CONCLUSIONS

WBCs have been utilized for years to inform cancer diagnosis and patient prognosis. Similarly, WBC derived gDNA has often been used in liquid biopsy as an intra-individual reference to differentiate somatic germline mutations from ctDNA mutations. This work combines these two established methods showing that intra-individual analysis can be used to provide additional relevant information to methylation cfDNA analysis to better define disease signal. IIX scoring was utilized to increase overall achieved specificity resulting in a PPV of up to 80.7%. Cell type deconvolution showed that one of the subjects identified as false positive by cfDNA correlated with memory T cell frequency, and read level comparisons were used to show that cancer patients had more abnormal reads than non-cancer patients.

## REFERENCES

1. Loyfer et al. A DNA methylation atlas of normal human cell types. Nature. 2023:613,355-364
2. Sales et al. Enhanced cell deconvolution of peripheral blood using DNA methylation for high-resolution immune profiling. Nat. Commun. 2022:13.761

## ACKNOWLEDGEMENTS

We gratefully acknowledge all individuals who participated in CORE-HH (NCT05435066), without their contributions this research would not have been possible.
Figure 2 created using Biorender.com

## DISCLOSURES

NCT05435066 and this analysis was sponsored by Harbinger Health. Cambridge, MA

www.harbinger-health.com