Detection of Somatic Copy Number Aberrations in cfDNA from Patients with Solid Tumor Malignancies

Neil Peterman, Alex Robertson, Abhik Shah, Rohith Srivas, Nicole Lambert, Tim Wilson, Peter George, Becky Wong, Ayse Tezcan, Ram Yalamanchili, John Spinosa, and Haluk Tezcan

Lexent Bio, San Diego, CA & San Francisco, CA

Highlights

- **Performed whole genome sequencing (WGS) on cfDNA extracted from plasma to measure CNAs**
- **Confident detection of CNAs in samples from 61% of late-stage cancer patients, with no CNAs detected among samples from 39 healthy participants**
- **High concordance between CNAs called in plasma and tumor tissue**
- **Detection limits depend on the pattern of CNAs present in individual tumor genomes**

Abstract

Profiling tumor-derived cell free DNA (cfDNA) captured from plasma is an attractive non-invasive approach to identify cancer at an earlier stage, determine actionable mutations for targeted therapies, detect residual disease, and monitor response to therapy. Large-scale somatic copy number aberrations (CNAs) are nearly universal in cancer, making them a promising feature of cfDNA to study tumor biology and changes that occur on therapy. A key requirement for broad clinical utility of this signal is reliable and accurate detection of CNAs in cfDNA, particularly at low-to-moderate tumor fraction (<5%). We developed an assay and analysis pipeline to identify CNAs in cfDNA using whole genome sequencing. We confirmed that CNAs identified from plasma were highly concordant with those identified by tumor tissue sequencing from a cohort of 29 cancer patients. We then assessed the performance of our test using plasma samples taken from another cohort of 57 participants with late-stage solid tumor malignancies and 39 healthy participants (control). Initial results showed detection of CNAs in 61% of samples from cancer patients, and none among control samples. Next, we performed in silico dilution analysis to objectively measure our detection limits. In order to account for tumor fraction as well as the extent and amplitude of CNAs, all of which are expected to impact detectability, we defined a new metric, the number of aberrant copies per genome. Across a diverse set of tumor genomes, we found this metric exhibits a consistent limit of detection. Our results indicate that CNAs in cfDNA have the potential to enable blood-based monitoring for solid-tumor malignancies, a subject we are investigating in ongoing studies.

Cohort

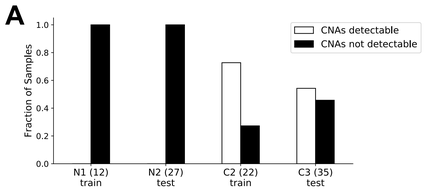
Tumor tissue samples were collected from cohort C1 during intent-to-cure surgery, along with plasma samples prior to surgery. Plasma samples from C2 and C3 were collected prior to starting a new line of therapy. Plasma samples were also collected from healthy participants, N1 and N2. We processed all samples with WGS, reaching a median coverage of 20X.

Cohort	**C1** train	**C2** train	**C3** test	**N1** train	**N2** test
N	29	22	35	12	27
Age					
Median	65	68	70	36	63
Range	39 - 81	38 - 89	30 - 87	21 - 61	27 - 73
Sex					
Female	21	9	29	3	7
Male	8	13	6	9	20
Cancer Stage					
I-III	29	0	1		
IV	0	22	34		
Cancer Type					
Breast	15	4	17		
Lung	0	7	16		
Colon	8	2	0		
Rectum	3	3	0		
Other	3	6	2		

Detection Sensitivity

We developed a pipeline to call CNAs based on read counts using a Hidden Markov Model (HMM). We set a stringent detection threshold to determine which samples have potentially actionable calls, requiring at least 3% of the genome to show confident CNA calls. This showed detection CNAs in the majority of samples from late-stage cancer patients (A). The test is highly specific with no CNAs detected among normal samples. Performance was similar in training and validation sets, with no significant difference between cohorts C2 and C3 (chi-squared test).

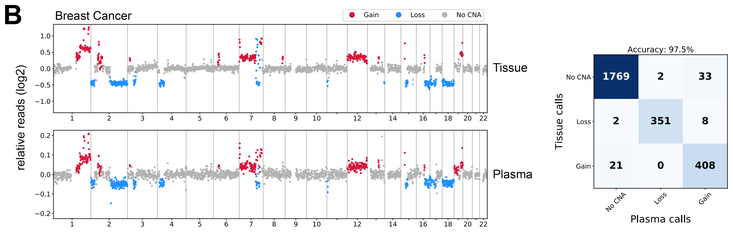
A



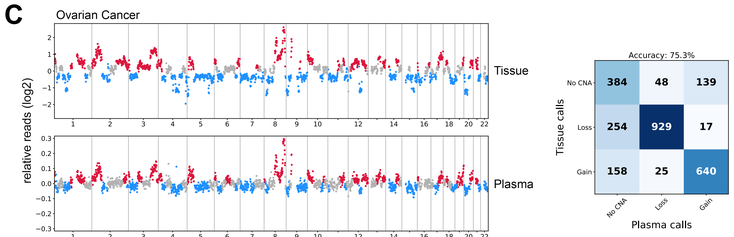
Concordance of Tissue and Plasma

3 plasma samples from cohort C1 had confident CNA calls. All 29 tissue samples indicated clear CNAs, however a lower detection rate in plasma is consistent with lower tumor fraction expected in earlier stage patients. We compared CNA calls from plasma and tissue for each of these patients. The left panels show corrected counts in 1Mb bins across the genome, as well as calls indicating gain (> 2 copies, red) or loss (< 2 copies, blue). Right panels show the number of bins with each pair of calls.

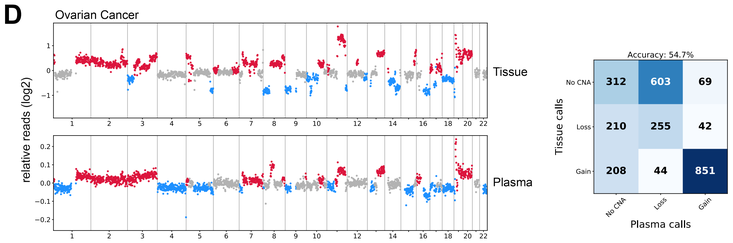
B



C



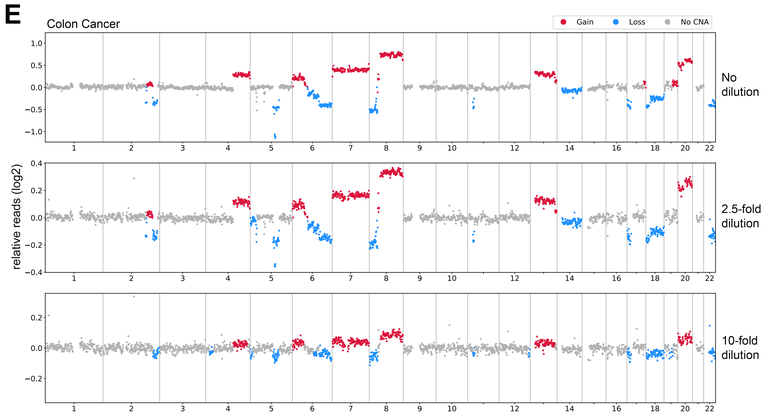
D



For two samples (B-C), we found high concordance with tissue. One sample (D) showed some agreement, but substantial disagreement as well; the majority of CNAs in tissue were identified in plasma, however plasma additionally showed many loss calls that are not present in the tissue sample. Discordances in samples with a large fraction of the genome copy number aberrant such as this one could be explained by ambiguity in which discrete segments are 2-copy. Additionally, there may be real clonal differences between tissue and plasma.

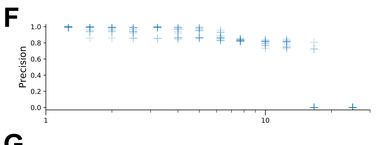
Limit of Detection Analysis

E

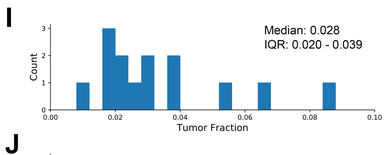


We carried out in silico dilutions of cancer samples using samples from our healthy cohorts in order to determine our detection limits for CNAs. CNA calls are similar for one sample from a metastatic colon cancer patient, even after a simulated 2.5-fold or 10-fold dilution (E-G). For each sample, we defined a critical dilution factor where less than 80% of simulated dilutions have CNAs above the detection threshold, for example 17X in the above example (H).

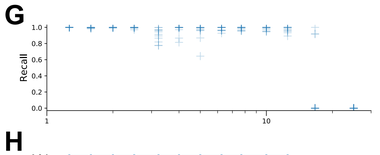
F



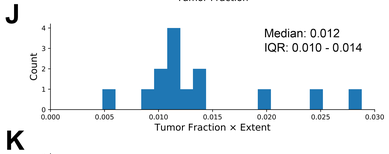
I



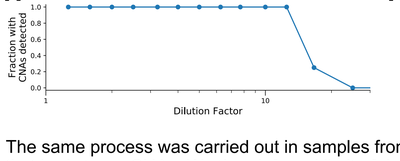
G



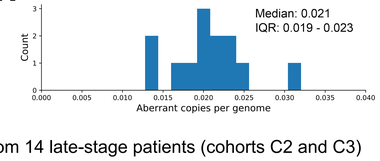
J



H



K



The same process was carried out in samples from 14 late-stage patients (cohorts C2 and C3) that had strong CNAs. We then inferred limit of detection metrics for each sample using the critical dilution factor:

- Tumor fraction (I) is a fitted parameter of the HMM used to call CNAs. Since cancer genomes with a larger extent of CNAs should be easier to detect, it is not surprising that tumor fraction alone is a poor metric for detection limits (I).
- The product of tumor fraction and the fraction of the genome with confidently called CNAs has a narrower distribution (J).
- A third metric, aberrant copies per genome, quantifies the average deviance from normal among copy number estimates from individual bins. This metric shows the most consistent limit of detection across the set of samples (K).

Acknowledgements

We would like to thank the patients and their families for participating in this study. We thank Dr. David Chan, Dr. James M. Sinclair, Dr. Andrew A. Davis, Dr. Wade T. Iams, Dr. Michael S. Oh, Dr. Robert W. Lentz, and Dr. Young Kwang Chae for help with recruitment and gathering medical data.