Tumor Shedding into Cell-free DNA (cfDNA) is Associated with High-mortality Cancers

2019 Cold Spring Harbor Laboratory Meeting:
The Biology of Genomes
May 7–11, 2019
Cold Spring Harbor, NY

Oliver Venn, DPhil; Earl Hubbell, PhD; Onur Sakarya, PhD; Christopher Chang, PhD; Meredith Halks-Miller, MD, FCAP; Kristan Steffen, PhD; Eric T. Fung, MD, PhD; Anne-Renee Hartman, MD; Tara Maddala, PhD; Alexander M. Aravanis, MD, PhD

GRAIL, Inc., Menlo Park, CA

INTRODUCTION

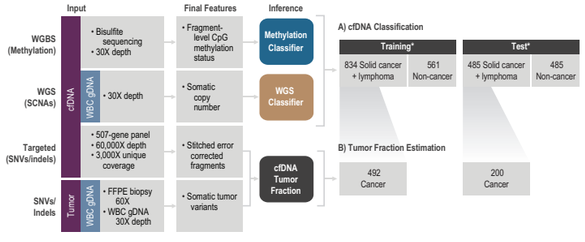
- Estimated cell-free DNA (cfDNA) tumor fraction (ECTF), the proportion of tumor molecules in a cfDNA sample, is a direct measurement of signal for cfDNA cancer applications.
- To determine the limit of detection (LOD) requirements across cancer types, there is a need for increased understanding about the nature of ECTF and factors influencing its abundance in cfDNA.
- Previous studies indicated that ECTF varies among cancer types and individuals.[1] However, these studies were limited by small sample sizes and surveyed few cancer types.
- This analysis leverages data from the Circulating Cell-free Genome Atlas (CCGA) study (NCT02889978; a prospective, multi-center, observational, case-control study) to examine ECTF across stages in >20 tumor types.

METHODS

Cancer Detection Classification on cfDNA

- Independent of tumor sequencing, cfDNA classifiers of cancer vs non-cancer for whole-genome bisulfite (WGBS) and whole genome (WGS) (both sequenced at 30x depth) were trained on a training dataset, locked, and then applied to a test dataset (Figure 1A).
- Of the cancers included in classifier training and evaluation, ECTFs were analyzed for 692 participants (Figure 1B), representing 22 cancer types (Table 1).
- cfDNA cancer detection classification performance for WGBS and WGS were assessed against ECTF.

Figure 1. Experimental Design



Of the 1319 solid cancer and lymphomas with cfDNA sequencing, 692 had FFPE biopsies and WBC whole genomes sequenced. Features identified with each cfDNA assay are indicated (WGBS: methylation; WGS: somatic copy number alterations). Somatic single nucleotide variants and insertions/deletions were called from matched tumor biopsy and WBC sequencing. Tumor biopsy variants enabled the verification of tumor signal in high depth targeted cfDNA across 22 different cancer types.

Table 1: FFPE Biopsy Samples

Cancer Type	N
Breast	323
Colorectal	54
Lung	53
Prostate	47
Uterine	31
Renal	30
Esophageal	20
Lymphoma	15
Remaining*	119

**Includes 14 head and neck, 13 ovarian, 13 pancreas, 13 thyroid, 11 bladder, 11 gastric, 10 melanoma, 8 cervical, 8 hepatobiliary, 7 anorectal, 6 cancers of unknown primary, two mesothelioma, two gastrointestinal stromal tumor, one anal.*

Tissue Sample Processing

- Formalin-fixed, paraffin embedded (FFPE) tumor biopsy samples were provided from CCGA clinical study sites.
- Tissue was macrodissected from unstained slides to yield a tumor nuclear volume between 0.5 to 1.5 mm^3 from a tissue region that was at least 80% tumor as assessed by review of adjacent H&E slides.
- Matched WBC and tumor samples were processed for DNA extraction and WGS (target sequencing depth 30x WBC and 60x biopsy).

Tumor Biopsy Variant Calling and cfDNA Mutant Allele Counting

- Tumor variants were called using WBC to control for germline variant contamination through the SENTIEON TNSeq caller.[2]
  - Variant calls were further refined using in-house filters to jointly optimize the recovery of putative true-positive somatic variants[3] while minimizing retention of known false positives identified from recurrent events in the cfDNA and tissue samples.
- Targeted sequencing was applied to matched cfDNA covering 507 genes and intergenic regions. Reads were collapsed and stitched to high accuracy fragment sequences using unique molecular identifiers and fragment positions (mean collapsed coverage: 3,000x).
- Given the filtered variant calls for a tumor sample, the corresponding collapsed cfDNA fragments were assessed for mutant allele support and the unique molecule depth recorded.

Statistical Modeling of Estimated cfDNA Tumor Fraction Based on Tumor Variants

- The ECTF for each sample was calculated using the fraction of mutant reads contributed from the tumor to the cfDNA sample.
- For each participant, we computed the likelihood of observing the cfDNA tumor mutant allele counts corresponding to the alleles found in the tumor for a given value of the ECTF.
- Allele counts were approximated as a Poisson distribution where the rate is controlled by local sequencing depth, ECTF, and allele frequency in the tumor.
- Individual variants were treated as independent for purposes of aggregating the total likelihood.
- We summarize the likelihood by reporting the median ECTF value per sample and include approximate confidence intervals from the same density.
- The observed sample-to-sample variation in ECTF levels was very large. Consequently small technical corrections were not implemented.

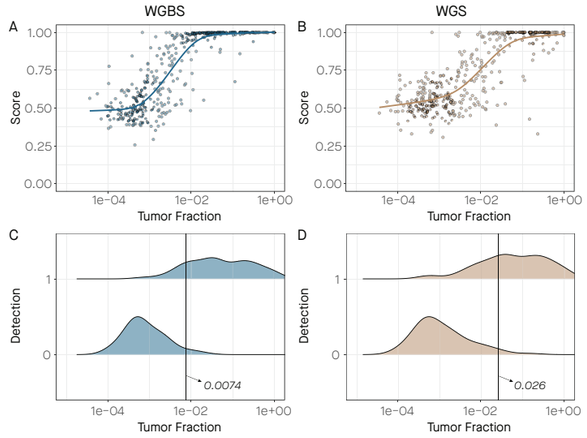
Cancer Type Mortality

- Incidence-weighted 5-year survival was obtained from Surveillance, Epidemiology, and End-Results (SEER [SEER18 ages 50+, 2000-2014]), and used to categorize surveyed cancer types into high mortality (bottom 10th percentile: esophageal, gastric, hepatobiliary, lung, pancreas) and low mortality (top 90th percentile: breast, prostate, thyroid).

RESULTS

- In targeted cfDNA sequencing, positive evidence of circulating tumor DNA (at least one alternative allele read, typical molecular depth 3000x) was observed in 420 of 692 participants. This population was used for all subsequent analyses.
- ECTF was an important covariate of classification performance, explaining >99% of the variance in WGBS and WGS classifier detection scores (Figure 2A and 2B).
- Classifiers' limits of detection (LOD, ECTF with 50% detection probability at 98% specificity) for WGBS and WGS were 0.73% (bootstrap [q5, q95], 0.41, 1.1; Figure 2C) and 2.7% (0.99, 5.9; Figure 2D).

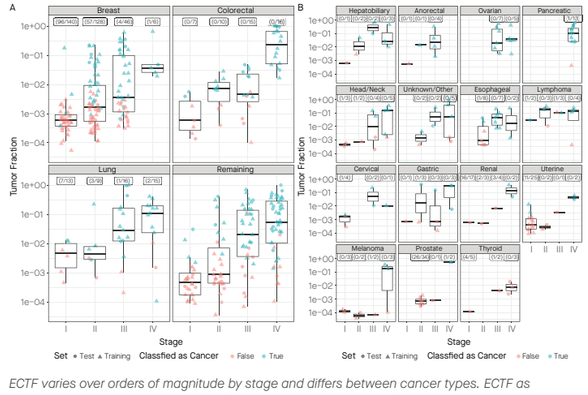
Figure 2. WGS and WGBS Classification Score Varies with Estimated cfDNA Tumor Fraction with Different Limit of Detection



The relationship between cfDNA cancer detection score and ECTF for A) WGBS and B) WGS assays. ECTFs (median of likelihood density) were calculated as the fraction of mutant reads contributed from the tumor to the cfDNA targeted panel sequencing. A monotone increasing p-spline was applied to WGS and WGBS scores on ECTFs across 113 participants and 18 cancer types in the test set for samples with positive read evidence. Dots represent estimates for each participant. ECTFs are presented on base 10 logarithmic scale. The empirical LOD (thick, annotated vertical line) was calculated by logistic regression of the cancer detection rate at 98% specificity against ECTF for C) WGBS and D) WGS classifiers for 121 CCGA substudy test set samples with positive read evidence. Densities for detected versus not detected are plotted. The WGBS classifier had a 4-fold lower LOD than WGS.

- ECTF increased with stage across cancer types (Figure 3A-B), however, ECTF varied by orders of magnitude within a given stage.
- Differences were observed between cancer types. For example, melanoma and prostate stage I-III cancers had small ECTF limiting cancer detection. In contrast, lymphoma was observed to have consistent, large ECTF across stages I through IV.

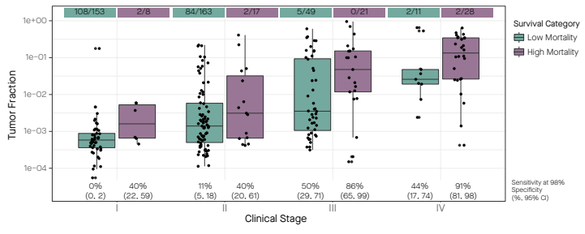
Figure 3. Distribution of Estimated cfDNA Tumor Fraction by Stage and Cancer Type



*ECTF varies over orders of magnitude by stage and differs between cancer types. ECTF as calculated by comparing targeted cfDNA sequencing with tumor WGS results (for samples with positive read evidence) was reported by stage for breast cancer, colorectal cancer, lung cancer, and other cancers in aggregate (A), and by each cancer type (B). Individual participant ECTFs are indicated by triangles (training set) and circles (test set), with symbol color indicating WGBS detection at 98% specificity (detected: blue; not detected: red). The proportion of samples for which no tumor variant evidence was observed in cfDNA is recorded for each distribution in beveled boxes. *Includes two neuroendocrine, two mesothelioma, two gastrointestinal stromal tumor, one anal, and four adenocarcinomas (not otherwise specified) of unknown primary origin.*

- High mortality cancers were observed to have larger ECTF within each stage than low mortality cancers with higher cfDNA classification sensitivity (Figure 4).
- Negative read evidence occurred more frequently at earlier stages.
  - For stages I and II, low- vs high-mortality cancers had greater rates of positive read evidence (192/316 vs 4/25, respectively; p=0.005, Fisher's exact test).

Figure 4. Within Stage, Deadlier Cancers have Higher Estimated cfDNA Tumor Fraction



Using incidence-weighted 5-year survival from Surveillance, Epidemiology, and End-Results (SEER), cancer types were categorized into high mortality (green, bottom 10th percentile: esophageal, gastric, hepatobiliary, lung, pancreas) and low mortality (purple, top 90th percentile: breast, prostate, thyroid). The distribution of ECTF for samples with positive read evidence within stage are summarized by boxplots; each participant ECTF recorded as a dot. The fraction of samples with no cfDNA read evidence for tumor variants are recorded in boxes (top). WGBS detection sensitivity at 98% specificity in the test set with 95% CIs are recorded (bottom). High mortality cancers showed larger ECTF than low mortality within stages II-IV (Kolmogorov-Smirnov test, p<0.05) and increased WGBS detection sensitivity within stages I-IV.

CONCLUSIONS

- Characterizing ECTF using the fraction of mutant reads contributed from the tumor to the cfDNA sample in high-depth sequencing indicated that:
  - ECTF was strongly associated with cfDNA classifier performance;
  - Tumor shedding combined with LOD of a given classifier and assay appeared to drive the observed differences in detection among cancer types;
  - Cancers with high mortality appeared to have higher ECTF.
- Together, these data suggest that cfDNA-based assays may detect high mortality cancers at earlier stages due to increased cfDNA shedding among those cancer types.

References

1. Abbosh C, Birkbak NJ, Wilson GA, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature.* 2017;545(7655):446-451.
2. Freed D et al. The Sentieon Genomics Tools - A fast and accurate solution to variant calling from next-generation sequence data. *bioRxiv.* January 2017:115717.
3. Chakravarty D, Gao J, Phillips SM, et al. OncoKB: A Precision Oncology Knowledge Base. *JCO precision oncology.* 2017;2017:10.1200/PO.17.00011.

Disclosures: All authors are employees of GRAIL, Inc. with stock/options to own stock in the company.