Many cell-free DNA (cfDNA) mutations are derived from clonal hematopoiesis: Implications for interpretation of liquid biopsy tests

Pedram Razavi[1], Bob T. Li[1], Chenlu Hou[7], Ronglai Shen[2], Oliver Venn[7], Raymond S. Lim[4], Earl Hubbell[7], Ino De Bruijn[4], Qinwen Liu[7], Ravi Vijaya Satya[7], Hui Xu[7], Ling Shen[7], Amy Sehnert[7], Tara Maddala[7], Michael F. Berger[1, 4, 5], Alex Aravanis[7], Jorge S. Reis-Filho[4], Mark Lee[7], Howard Scher[1, 6], David B. Solit[1, 3, 5], Jose Baselga[1, 3]

1. Department of Medicine, Memorial Sloan Kettering Cancer Center (MSKCC), 1275 York Avenue, New York, NY, USA. 2. Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center. 3. Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center. 4. Department of Pathology, Memorial Sloan Kettering Cancer Center. 5. Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center. 6. Genitourinary Oncology Service, Department of Medicine, Memorial Sloan Kettering Cancer Center. 7. GRAIL, Menlo Park, CA 94402 USA.

Background

- A large fraction of cell-free DNA (cfDNA) fragments are known to be derived from hematopoietic cells (Lui 2002).
- Sequencing of genomic DNA (gDNA) from white blood cells (WBC) in circulation has revealed an age-dependent presence of clones carrying somatic variants—a phenomenon now termed clonal hematopoiesis (Genovese 2014).
- Somatic alterations in cfDNA can be tumor-derived but also could represent somatic changes associated with clonal hematopoiesis.
- We applied a high-intensity sequencing approach (ultra-deep sequencing with broad genomic coverage) to characterize both cfDNA and WBC gDNA from cancer and non-cancer subjects.

Objective

- To evaluate the contribution of clonal hematopoiesis to the variants observed in cfDNA through high-intensity sequencing of both plasma cfDNA and matched WBC gDNA.

Methods: Patient Populations

Cancer Patient Population:

- Metastatic breast, lung, or prostate cancer, either de novo or with progressive disease on current therapy.
- All patients have provided written informed consent to an MSK institutional protocol (NCT01775072) allowing research of cfDNA and clinical tumor sequencing.
- Two tubes of blood collected in Streck Cell-Free DNA BCT.

Healthy Donor Population:

- Enrolled through San Diego Blood Bank.
- Self reported as non-smokers and cancer-free.
- Two tubes of blood collected in Streck Cell-Free DNA BCT.

Methods: Analysis

cfDNA and WBC gDNA variant calling pipeline

- The variant calling pipeline includes the following steps:
  - Read alignment, error correction (consisting of read collapsing and stitching of mate pairs), *de novo* assembly, variant calling, and variant filtering.
  - Variants are filtered using two approaches (1) heuristics applied based on the surrounding sequence context and type and quality of reads supporting the variant and (2) empirical noise levels observed in a set of healthy samples:
    - Blood samples from an independent cohort of 24 non-cancer (self-reported) individuals (vendor sourced), were used for noise modeling for the 151 cancer patients in this study.
    - The 47 healthy subjects were processed with a slightly different assay protocol and UMI set and therefore had 20 additional non-cancer subjects used to establish a noise model for that assay protocol.
  - Variants are further filtered based on dbSNP presence and restricted to protein coding variants.

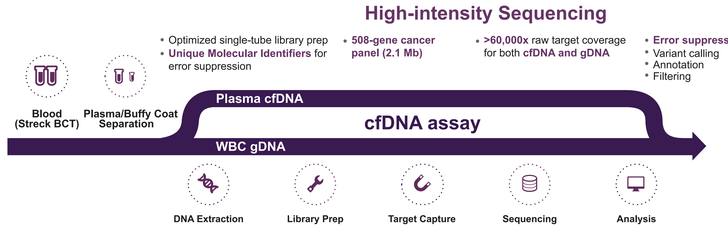
Quality filtering on cfDNA and WBC gDNA variant calls

- Analyses focus on matched variants in cfDNA and WBC gDNA meeting the following criteria:
  - Non-synonymous.
  - Allele frequency (AF) >= 0.1% and a minimum of 3 supporting reads.
  - In order to exclude germline, AF < 15% in cfDNA.
  - In order to exclude germline, AF < 35% in WBC gDNA.
  - Indels not in homopolymer >= 3.
  - Variants must be found in <=10 samples to avoid recurrent artifacts.
  - Unique coverage >=500 to exclude low coverage regions.

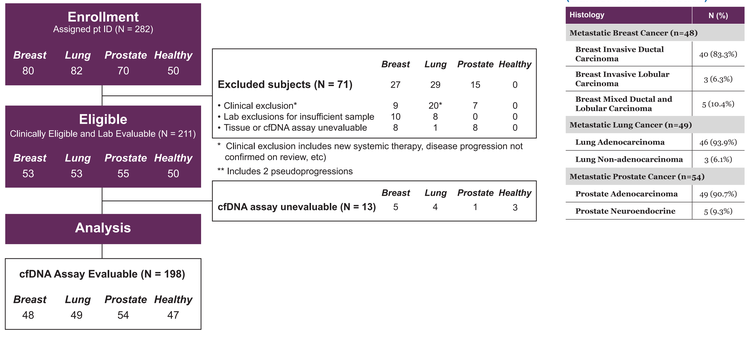
Statistical analysis of variants by age

- A Poisson regression model was used for analysis of the association between age, cohort, and mutation count (dependent variable).

Sample Workflows and Assays



Patient Disposition



	Breast	Lung	Prostate	Healthy
Excluded subjects (N = 71)	27	29	15	0
• Clinical exclusion*	9	20*	7	0
• Lab exclusions for insufficient sample	10	8	0	0
• Tissue or cfDNA assay unevaluable	8	1	8	0

* Clinical exclusion includes new systemic therapy, disease progression not confirmed on review, etc)

** Includes 2 pseudoprogressions

	Breast	Lung	Prostate	Healthy
cfDNA assay unevaluable (N = 13)	5	4	1	3

Patient Characteristics (Cancer Characteristics)

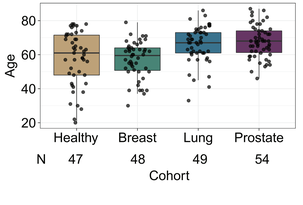
Histology	N (%)
Metastatic Breast Cancer (n=48)	
Breast Invasive Ductal Carcinoma	40 (83.3%)
Breast Invasive Lobular Carcinoma	3 (6.3%)
Breast Mixed Ductal and Lobular Carcinoma	5 (10.4%)
Metastatic Lung Cancer (n=49)	
Lung Adenocarcinoma	46 (93.9%)
Lung Non-adenocarcinoma	3 (6.1%)
Metastatic Prostate Cancer (n=54)	
Prostate Adenocarcinoma	49 (90.7%)
Prostate Neuroendocrine	5 (9.3%)

Patient Characteristics

Patient Characteristics	Breast (n=48)	Lung (n=49)	Prostate (n=54)	Healthy (n=47)
Age at enrollment				
Mean (SD)	57.0 (10.76)	66.1 (10.86)	67.6 (9.55)	57.8 (16.03)
Median	60	67	68	61
Range	30, 79	33, 86	46, 87	20, 78
Gender, N (%)				
Female	48 (100.0%)	32 (65.3%)	N/A	24 (51.1%)
# of lines of therapy, N (%)				
0	22 (45.8%)	28 (57.1%)	18 (33.3%)	
1	2 (4.2%)	12 (24.5%)	15 (27.8%)	
2	5 (10.4%)	4 (8.2%)	9 (16.7%)	
>=3	19 (39.6%)	5 (10.2%)	12 (22.2%)	
Receptor Status, N (%)				
HR+/HER2+	3 (6.3%)	N/A	N/A	
HR+/HER2-	34 (70.8%)	N/A	N/A	
HR-/HER2+	3 (6.3%)	N/A	N/A	
Triple Negative	8 (16.7%)	N/A	N/A	

Distribution of Age by Cohort

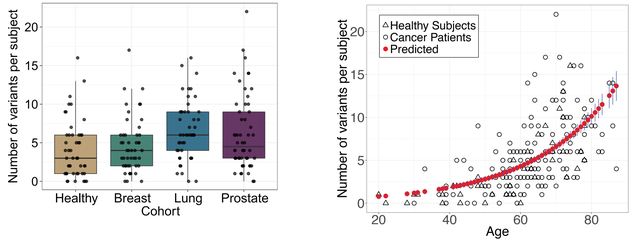
Healthy subjects and breast cancer patients tend to be younger than lung cancer and prostate cancer patients



Matched Variants in cfDNA and WBC Are Associated with Age

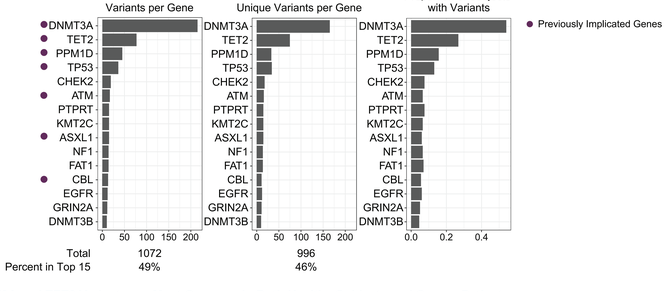
- Positive association with age: p-val <0.0001.
- The association of number of matched variants with age does not appear to be different between healthy subjects and those with breast/lung/prostate cancer in this cohort. Interaction of age with cancer not significant: p = 0.08.

Matched Variants in cfDNA and WBC Per Subject
(after Quality Filter)

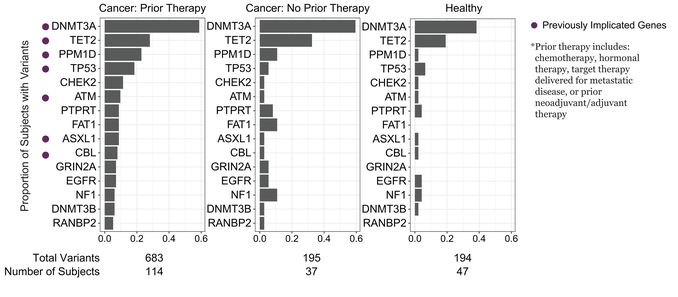


Variants Found in Both WBC and cfDNA are Preferentially in Previously Described Clonal Hematopoiesis Genes

Variants are in genes previously associated with clonal hematopoiesis (Genovese 2014).



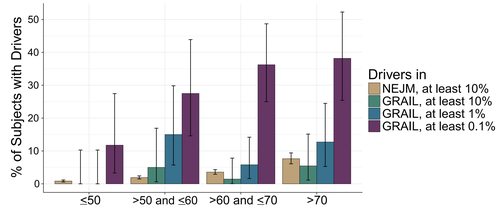
DNMT3A and TET2 Variants are Most Common in Both Healthy Subjects and Cancer Patients



*Prior therapy includes: chemotherapy, hormonal therapy, target therapy delivered for metastatic disease, or prior neoadjuvant/adjuvant therapy

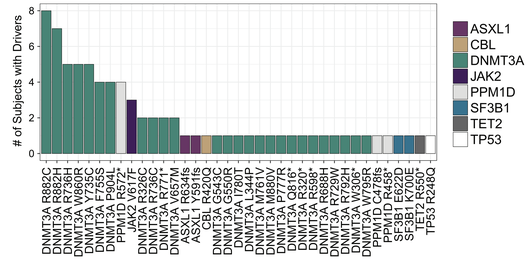
Improved LOD Increases Detection of Clonal Hematopoiesis Variants

- NEJM 2014 study (Genovese 2014) identified 217 unique candidate drivers in WES.
  - Disruptive (nonsense, frameshift, splice-site) mutations in *DNMT3A*, *TET2*, *ASXL1* and *PPM1D*; missense in *DNMT3A*.
  - Recurrent (≥ 7 patients) in hematopoietic and lymphoid cancers in COSMIC.
- 212 of 217 previously identified drivers are covered by cfDNA panel.
- 37 unique drivers (77 events) were detected in quality filtered WBC variants (total = 1072) among 198 subjects.



WBC Variants Contain Previously Identified Drivers

- Matched variants in cfDNA and WBC contain 77 driver variants (37 unique variants).
  - Drivers supported by a median of 12 unique reads in WBC gDNA (3 - 533).
  - Drivers supported by a median of 18 unique reads in cfDNA (3 - 617).



Summary and Conclusions

- This high-intensity sequencing approach applied to both plasma cfDNA and WBC gDNA reveals that many cfDNA variants arise from clonal hematopoiesis:
  - The most common variants matched between cfDNA and WBC gDNA are in genes previously implicated in clonal hematopoiesis, including *DNMT3A*, *TET2*, *PPM1D*, and *TP53*.
  - As with previous WBC gDNA sequencing studies, the number of cfDNA variants attributable to clonal hematopoiesis increases with age.
- The limit of detection of our sequencing approach suggests that clonal hematopoiesis variants may be present at low allele frequencies and possibly prevalent in more patients than previously reported. This hypothesis will require further evaluation.
- Accurate interpretation of cell-free DNA assays for detection of circulating tumor DNA must account for the confounding signals arising from clonal hematopoiesis, and may require deep sequencing of WBC gDNA (Abstract #11528, Poster #228, Razavi et al).

References

Lui, Y.Y., Chik, K.W., Chiu, R.W., Ho, C.Y., Lam, C.W., Lo, Y.M. (2002). Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem, 48*, 421-7.

Genovese, G., Kähler, AK., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakoum, S.F., Chambert, K., Mick, E., Neale, B.M., ... McCarroll, S.A. (2014). Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N Engl J Med, 371*, 2477-87. https://doi.org/10.1056/NEJMoa1409405