Distinguishing Between Cancer-Related Mutations and Clonal Hematopoiesis Using Cell-Free RNA (cfRNA) Expression Levels in Machine Learning Model

Maher Albitar, MD[1]; Hong Zhang, PhD[1]; Andrew Ip, MD[2]; Jeffrey Estella, BSc[1]; Ahmad Charifa, MD[1]; Wanlong Ma, BSc[1]; Arash Mohtashamian, MD[1]; Martin Gutierrez MD[2]; Andrew Pecora, MD[2] and Andre Goy MD[2]

1 Genomic Testing Cooperative, Irvine, CA, 2 John Theurer Cancer Center at Hackensack University Medical Center, Hackensack, NJ

#3043

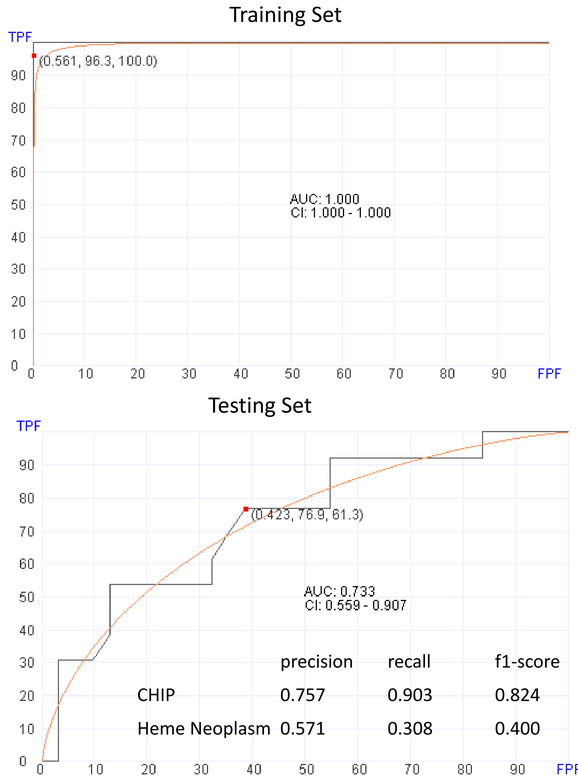
Introduction

Liquid biopsy is currently considered an important part of the clinical practice of oncology. However, proper interpretation of somatic mutations detected in liquid biopsy remains a challenge. Distinguishing cancer-related mutations from mutations resulting from clonal hematopoiesis of indeterminate potential (CHIP) can be difficult and a source of misinterpretation of liquid biopsy findings. Mutations in ASXL1, TET2 and DNMT3A genes, which are the whole mark of CHIP can be seen 5-10% of solid tumors. this also makes distinguishing between mutations in circulation resulting from solid tumors vs CHIP more difficult. We explored using cell-free RNA (cfRNA) profiling as a mean for distinguishing between mutations detected as CHIP vs mutations detected as cancer.

Methods and Materials

cfDNA and cfRNA from 102 patients with confirmed solid tumors, 93 patients with hematologic neoplasms and 40 patients with CHIP abnormalities were sequenced using a panel of 1501 gene for RNA and 284 genes for DNA. The solid tumors included lung, breast, ovarian, and colorectal. The hematologic neoplasms included lymphoid and myeloid neoplasms. Using a machine learning algorithm, we first selected the relative genes that distinguish between two classes using two criteria: performance of each gene with K-fold cross-validation (K=12) and second based on stability measure using statistical significance tests. The selected genes were used to predict one class from the other using Random Forest or Extreme Gradient Boosting (XGBoost). For training, we used 67% of cases and for testing we used 33% of cases.

Mutations associated with hematologic neoplasms can be distinguished from CHIP using cfRNA from 40 genes.



Top 10 genes and relevant pathways involved in this algorithm:

CAV1, PBX1, CDK8, PRKACA, SMARCA5, SRC, DOT1L, SUZ12, CD14, JAK1,

Pathways

MAPK signaling: CD14, MEAF6, PDGFA, and NF2

Th1 and Th2 cell differentiation: NF2, IL1RAP, and JAK1

PI3K-AKT: JAK1, PDGFA, and IL1RAP.

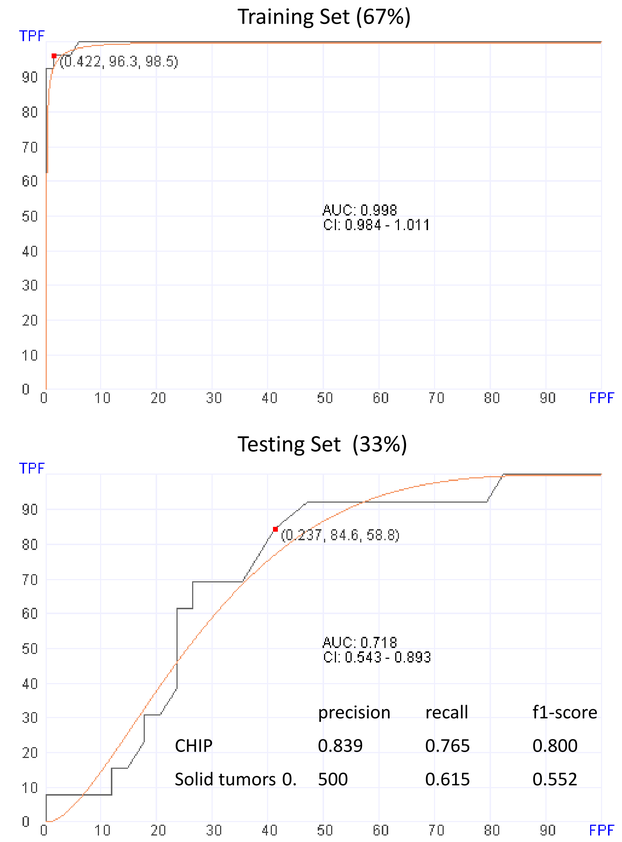
Conclusions:

-Analyzing cfRNA expression levels adds another level of confidence to the ability of interpreting mutations detected in liquid biopsy testing.

-Machine learning algorithms are needed for interpreting low level cfRNA biomarkers in circulation

-cfRNA in machine learning when combined with cfDNA is particularly important in liquid biopsy testing for minimal residual disease (MRD) when a tissue baseline sample is not tested (tumor-agnostic).

Cancer-related mutations can be distinguished from CHIP using cfRNA levels of 25 genes.



Top 10 genes and relevant pathways involved in this algorithm:

BRAF, HOOK3, EXOSC6, CD14, GNAI1, EIF4E, HNF1A, MAFB, H2AX, YY1AP1.

Pathways:

Transcription misregulation: HOOK3, CD14, YY1AP1, ERCC6, RUNX1

Phagozome: CD14, HIPK2, CD34, FCGR2B