Fragmentomics of cell free DNA from targeted panels in genitourinary malignancies

Amy K. Taylor[1,2], Kyle T. Helzer[3], Marina N. Sharifi[1,2], Jamie M. Sperger[1], Yue Shi[3], Matti Annala[4], Shannon R. Reese[1], Katie R. Kaufmann[1], Jennifer L. Schehr[1], Nan Sethakorn[5], David Kosoff[1,6], Christos Kyriakopoulos[1,2], Andrew J. Armstrong[7], Rana R. McKay[8], Felix Y. Feng[9], Kari B. Wisinski[1,2], Hamid Emamekhoo[1,2], Alexander W. Wyatt[4], Joshua M. Lang[1,2] & Shuang G. Zhao[1,3]

Affiliations: 1 Carbone Cancer Center, University of Wisconsin, Madison, WI, USA; 2 Department of Medicine, University of Wisconsin, Madison, WI, USA; 3Department of Human Oncology, University of Wisconsin, Madison, WI, USA; 4 Vancouver Prostate Centre, Department of Urologic Sciences, University of British Columbia, Vancouver, British Columbia, Canada; 5 Department of Medical Oncology, Loyola University, Chicago, Illinois, USA; 6 William S. Middleton Memorial Veterans Hospital, Madison, Wisconsin, USA; 7 Duke Cancer Institute Center for Prostate and Urologic Cancers, Department of Medicine, Duke University, Durham, North Carolina, USA; 8 Moores Cancer Center, University of California, San Diego, La Jolla, California, USA; 9 Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA, USA

Background: The detection of genomic alterations in cancer is critical for identifying clinically actionable alterations for treatment decisions. Tumor samples historically have been required, but obtaining tissue for molecular profiling is not always feasible, especially in the metastatic setting. The isolation and analysis of cell-free DNA (cfDNA) including circulating tumor DNA (ctDNA) via blood-based "liquid" biopsies offers non-invasive sampling of the tumor. Recently, fragmentation patterns of cfDNA (i.e. "fragmentomics") have emerged as a method for inferring epigenomic and transcriptomic information. However, these analyses use whole-genome sequencing, which lacks the necessary depth to cost-effectively assess genomic alterations, limiting the application of these techniques clinically.

Methods: We developed a novel cfDNA fragmentomics machine learning approach for standard targeted cancer gene panels in order to identify genitourinary (GU) and non-GU cancers from two independent metastatic cancer cohorts: a published cohort from GRAIL (prostate cancer, non-GU cancers, normal samples, N=198) and an expanded institutional cohort from the University of Wisconsin (prostate adenocarcinoma, neuroendocrine prostate cancer (NEPC), renal cell carcinoma (RCC), bladder cancer, non-GU cancers, normal samples, n = 431).

Results: In the GRAIL cohort, 10-fold cross-validation AUCs were 0.987 for identifying prostate cancer, 1.00 for normal samples, and ranged from 0.922-0.958 for the non-GU cancers. In the UW cohort, 10-fold cross-validation AUCs were 0.950 for bladder cancer, 0.982 for prostate cancer, 0.993 for NEPC, 0.925 for RCC, and 0.980 for normal samples, and ranged from 0.874-0.954 for non-GU cancers. Our assay can sensitively detect and accurately distinguish GU and non-GU malignancies despite a median ctDNA fraction of only 0.076 in the GRAIL cohort, and a median ctDNA fraction of only 0.022 in the UW cohort.

Conclusions: Our enhanced machine learning approach for examining fragmentomics in standard cancer gene cfDNA panels unlocks the potential for these established assays to be utilized to answer critical clinical questions beyond somatic variant identification. The excellent performance of our innovative approach even in samples with low ctDNA fractions suggest potential applications such as multi-cancer early detection and minimal residual disease monitoring. This framework could dramatically expand the potential of already existing clinically used assays and minimizes barriers for continued biomarker development.

Funding Acknowledgments: R01 CA256157/CA/NCI NIH HHS/United States; P30 CA014520/CA/NCI NIH HHS/United States**;** R01 CA247479/CA/NCI NIH HHS/United States**;** UH2 CA260389/CA/NCI NIH HHS/United States**;** R37 CA241486/CA/NCI NIH HHS/United States

COI Disclosures: KTH has a family member who is an employee of Epic Systems. SGZ reports unrelated patents licensed to Veracyte, and that a family member is an employee of Artera and holds stock in Exact Sciences. KTH, SGZ, and the University of Wisconsin have filed a provisional patent on the work herein. FYF reports personal fees from Janssen Oncology, Bayer, PFS Genomics, Myovant Sciences, Roivant Sciences, Astellas Pharma, Foundation Medicine, Varian, Bristol Myers Squibb (BMS), Exact Sciences, BlueStar Genomics, Novartis, and Tempus; other support from Serimmune and Artera outside the submitted work. Integrated DNA Technologies (IDT, Coralville, IA) assisted in a pilot project to assess the performance characteristics of the UW panel before purchase, but played no role in this study. All other authors have declared no conflicts of interest. Data Sharing Raw sequencing data from the GRAIL dataset are available at the European Genome Archive (Dataset ID EGAD00001005302). Our institutional protocol does not allow unrestricted public access to the raw sequencing data. Therefore, data sharing requests must be submitted to the University of Wisconsin-Madison for approval. For samples from the two clinical trials (NCT03090165, NCT03725761), these trials are still ongoing, and data sharing requests must be submitted to the trial organizers.