A computational approach to identifying ancestry related alternatively spliced and differentially expressed genes in African American and White prostate cancer

Muthana Al Abo[1], Steven R. Patierno[1,2] and Jennifer A. Freedman[1,2]

[1] Duke Cancer Institute, Duke University Medical Center, Durham, NC, 27710, USA
[2] Department of Medicine, Division of Medical Oncology, Duke University Medical Center, Durham, NC, 27710, USA

Prostate cancer (PCa) affects disproportionally men from differet population groups. The SEER's 2018 report reveals that PCa incidence and mortality rates are ~2 times higher among African American (AA) men in comparison with white men. In addition to differences in social, lifestyle and structural determinants of health, there is accumulating evidence for a biological contribution to racial/ethnic disparity in PCa. To date, most work focused on understanding further the molecular mechanisms underlying racial/ethnic disparity in PCa analyzes differential aggregate gene expression and mutation among cancer patients of different population groups. Our recently published work reported alternative splicing (AS) as an additional critical mechanism underlying prostate cancer aggressiveness and drug response in AA patients. Here, we analyzed the Genomic Data Commons for differential aggregate gene expression (2-fold mean change, $p < 0.001$, Wilcoxon rank sum test) and TCGASpliceSeq to analyze The Cancer Genome Atlas (TCGA) for AS (20% median change, percent spliced in) between PCa specimens from AA and white patients. From our analysis of the 307 PCa specimens from white patients and 49 PCa specimens from AA patients in TCGA, we identified 71 differentially expressed genes (DEGs) and 73 alternative splicing events (ASs) between the specimens from AA and white PCa patients. 51 of the DEGs (~72%) exhibit increased expression levels in PCa from AA patients compared with white patients. Notably, the genes that exhibit differential aggregate gene expression and the genes that exhibit AS do not overlap, indicating that ancestry-related aggregate gene expression and AS can be independent events. Among the AS events, the majority involve exon skipping (35 events, ~48%). The distribution of the remaining AS events includes alternative acceptor (9 events, ~12%), alternate terminator (8 events, ~10%), alternate donor (6 events, ~8%), alternate acceptor (4 events, ~5%) and mutually exclusive exon (2 events, ~1%). Clustering and pathway analysis of DEGs and ASs reveal that many of these genes function in pathways relevant to cancer development and progression, such as programmed cell death, DNA repair, signal transduction, gene expression and metabolism. These analyses increase understanding of molecular mechanisms underlying racial/ethnic disparity in PCa. Upon further study of

the function of these variants, such DEGs and ASs have the potential to become candidates for development of new precision medicine interventions.

Conflict of Interest: None.

Funding: This work was supported by a NIH Basic Research in Cancer Health Disparities R01 Award R01CA220314 to SRP PI and JAF Co-I.