DR. OLIVIA DENISE LARA (Orcid ID : 0000-0002-3540-5804)

DR. YUEXIN LIU (Orcid ID : 0000-0002-0828-4656)

Article type : Original Article

Pan-cancer clinical and molecular analysis of racial disparities

Short Title: Analysis of racial disparities

Olivia D. Lara, M.D.[1]*, Ying Wang, Ph.D. [2]*, Amma Asare, M.D.[3], Tao Xu, Ph.D.[4], Hau-Sheng Chiu, Ph.D.[5], Yuexin Liu, Ph.D.[2], Wei Hu, M.D.[1], Pavel Sumazin, Ph.D.[5], Shitanshu Uppal, M.D. [4], Lin Zhang, Ph.D.[6], J. Alejandro Rauh-Hain, M.D.[1], Anil K. Sood, M.D.[1,7]

Authors' Affiliations:

[1] Departments of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas

[2] Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas

[3] Baylor College of Medicine, Houston, Texas

[4] Department of Gynecologic Oncology, University of Michigan, Ann Arbor, Michigan

[5] Texas Children's Cancer Center, Houston, Texas

[6] Department of Obstetrics and Gynecology, University of Pennsylvania, Philadelphia

[7] Center for RNA Interference and Non-Coding RNA, The University of Texas MD Anderson Cancer Center, Houston, Texas
*Authors contributed equally to this work

Corresponding Author:
Address reprint requests to Dr. Anil K. Sood at the Department of Gynecologic Oncology and Reproductive Medicine, Unit 1362, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030 or at asood@mdanderson.org, 713-792-6810

Conflict of interest:
A.K.S. is an advisor for KIYATEC and Merck & Co., Inc., and a Bio-Path Holdings stockholder and has received research funding from M-Trap. The remaining authors declare no competing interests.

Funding:
O.D.L. is supported by a National Institutes of Health institutional training grant (5T32CA009599). This work was supported in part by other National Institutes of Health grants (P30CA016672, CA213759, P50CA217685, P50CA098258, and R35CA209904), the Blanton-Davis Ovarian Cancer Research Program, the American Cancer Society Research Professor Award, and the Frank T. McGraw Memorial Chair in Cancer Research (to A.K.S.).

Author Contributions
O.D.L., Y.W., A.R.H., and A.K.S. contributed to the study design, acquired and analyzed data, generated figures, and wrote the manuscript. A.A. analyzed data and generated figures. W.H. provided intellectual input. L.Z. provided intellectual input and established TCGAA. T.X., H.S.C., Y.L., S.U., and P.S. contributed to data acquisition and statistical

analysis. All authors contributed to the interpretation of data, vouched for the data analysis, contributed to the writing of the manuscript, and agreed to publication of this study.

Acknowledgements

We acknowledge The University of Texas MD Anderson Cancer Center Department of Scientific Publications for reviewing and editing of this manuscript.

Precis

AA race is associated with a survival disadvantage across two large cancer registries and highlights molecular differences in AA tumors which may contribute to differences in patient outcomes. This pan-cancer clinical and molecular analysis of AA race may help to identify actionable targets which could lead to new approaches to overcome cancer disparities.

Abstract

Background: Racial disparities in cancer outcomes are increasingly recognized, but comprehensive analyses, including molecular studies, are limited. Our objective in the present study was to perform a pan-cancer clinical and epigenetic molecular analysis of outcomes in African American (AA) and European American (EA) patients.

Methods: Cross-platform analyses using cancer databases (Surveillance, Epidemiology, and End Results Program database [SEER] and National Cancer Database [NCDB]) and a molecular database (The Cancer Genome Ancestry Atlas [TCGAA]) were

performed to evaluate clinical and epigenetic molecular differences between AA and EA patients based on genetic ancestry.

Results: In the primary pan-cancer survival analysis utilizing SEER database (2,045,839 patients [87·5% EA and 12·5% AA]), AA patients had higher mortality rates for 28 of 42 cancer types analyzed (hazard ratio [HR]>1·0). AA's continued to have higher mortality in 13 cancer types after adjustment for socioeconomic variables utilizing NCDB database (5,150,023 patients [11·6% AA and 88·4% EA]). We then analyzed molecular features of 5283 tumors in patients with available genetic ancestry data (87·2% EA and 12·8% AA). We identified genes with altered DNA methylation, along with increased microRNA expression levels unique to AA patients that are associated with cancer drug resistance. Increased miRNAs (*miR-15a, miR-17, miR-130-3p, miR-181a*) were noted in common among AAs with breast, kidney, thyroid or prostate carcinomas.
Conclusions: These results identified epigenetic features in AA cancer patients that may contribute to higher mortality rates compared to EA cancer patients. Therefore, a focus on molecular signatures unique to AAs may identify actionable molecular abnormalities.

Keywords: Pan-cancer, Healthcare disparities, African Americans, epigenomics, microRNAs, RNA, Long Noncoding

Total number:
Text (pages, abstract, main text, references, legends): word count 5266
Tables: 2
Figures: 2
Supporting files: supplementary appendix

Introduction

For many cancer types, authors have documented racial disparity in patient survival. In particular, African American (AA) patients have higher mortality rates than do all other racial groups for many cancer types.[1-3] For example, AA women are 40% more likely than European American (EA) women to die of breast cancer, whereas AA men are twice as likely as EA men to die of prostate cancer.[4] Many factors contribute to disparities in cancer outcomes, including socioeconomic factors, culture, diet, stress, and the macroenvironment.[5-7] Research into health disparities has primarily focused on these social determinants and their impact on health. However, other observations suggest a role for molecular differences in cancer survival disparities.[8] Compared with EA women, AA women have a twofold higher incidence of the inherently aggressive triple-negative breast cancer (TNBC).[9] Comparison of TNBC in AA versus EA patients has demonstrated a gene expression signature consistent with increased loss of BRCA1 expression, increased activation of insulin-like growth factor 1 receptor, and increased expression of vascular endothelial growth factor-activated genes in AA patients.[10] Additionally, AA patients with breast, head and neck, or endometrial cancer have higher levels of chromosomal instability and frequency of TP53 mutations and CCNE1 amplification than do white patients.[11]

Whereas some studies have explored somatic alterations in individual cancers,[12-14] little is known about how the macroenvironment affects the epigenetic landscape among the different races.[15,16] Epigenetic differences are particularly relevant because they can be shaped by environmental factors such as chronic stress, social interactions, and toxins.[15] Theses epigenetic changes can function as a liaison between social, cultural, environment factors and the genome.[17] Furthermore, race is a social construct and is often a proxy for chronic stressors that minority populations face, such as low

socioeconomic status, poor access to health care, dangerous environments, and interpersonal discrimination.[18] These unique stressors that a racial group experiences may impact their epigenome and ability to respond to disease, leading to poor outcome and response to therapy.

In this study, we used data from two large cancer registries to perform a pan-cancer mortality analysis of AAs compared with EAs. We also evaluated differences in epigenetic modifications, including non-coding RNA (microRNA and long-non coding RNA) expression and DNA methylation, in the two groups. We believe epigenetic modifications unique to AAs may contribute to differences in clinical outcomes.

Methods

Data Sources

In our primary survival analysis, data from the Surveillance, Epidemiology, and End Results (SEER) Program version 9 database, which contains information on cancer incidence and survival from 17 population-based cancer registries covering about 28% of the U.S. population, were used. Because the SEER database lacks comprehensive sociodemographic information, a secondary survival analysis using data from the National Cancer Database (NCDB), another U.S. cancer registry, was performed. The NCDB is a nationwide oncology outcomes database that includes information on about 70% of all invasive incident cancers diagnosed in the United States. Tumor registrars at participating hospitals document patient, tumor, and treatment characteristics (including information about initial surgery, chemotherapy, and radiotherapy) in addition to survival. The NCDB population consists of patients who received cancer care (treatment or diagnosis) from among 1400 cancer programs accredited by the Commission on Cancer. These survival data were paired with molecular data from The Cancer Genome Atlas (TCGA).[19] To define the AA and EA patient groups, the Cancer Genome Ancestry Atlas (TCGAA)[11] data were used. TCGA contains data such as clinical information, histopathology, and molecular information derived from information on samples obtained from more than 11,000 patients. TCGA microRNA (miRNA) and long non-coding RNA (lncRNA) sequencing data used in the present study were analyzed on the

Illumina HiSeq platform, whereas TCGA DNA methylation data were analyzed on the Illumina HM450k platform.

Study Population, Covariates, and Statistical Analyses

All cancer cases documented from 2000 to 2015 in the SEER 9 registry research data were identified. Only patients classified as EA or AA were included in our study. Additionally, patients with a prior history of cancer, with a survival or follow-up time shorter than 1 month, who were older than 100 years at diagnosis, who had non-cancer causes of death, who had tumor types found in fewer than 200 cases in the registry, or who had missing age or survival data were excluded. The patients had a total of 42 primary tumor types. The primary outcome was overall survival in AA and EA patients (see the Supplemental Appendix).

For cancer types identified in the SEER survival analyses with hazard ratios (HRs) for mortality greater than 1·0 in AAs compared with EAs, a secondary survival analysis was performed for cancer types in the NCDB. Cohort defined in supplemental appendix. Only tumor types with available TCGA data on molecular correlations were included in this analysis (Fig. S1 in the Supplementary Appendix).

Patients with 13 primary tumor types having a persistent survival disadvantage according to the NCDB survival analysis were subjected to molecular analyses. Based on the genetic ancestry information from TCGAA, EA and AA patients in TCGA were included in our study. This differed from SEER and NCDB which rely on self-identified race to identify AA and EA groups. The number of AA patients with skin cutaneous melanoma (SKCM) in TCGA was insufficient for testing, so this group was excluded from all analyses. The distribution of EA and AA patients in TCGA according to tumor type is shown in Table S1 in the Supplementary Appendix. For each TCGA tumor type of interest, feature-by-feature Wilcoxon tests were used to compare DNA methylation profiles, and feature-by-feature *t*-tests were run to compare the miRNA sequencing profiles for the AA and EA groups. A beta-uniform mixture model [20] was used to adjust

for multiple comparisons and estimate significant features of AA race at different false-discovery rates (FDRs). Long non-coding RNA (lncRNA) expression profiles were normalized using The Atlas of non-coding RNA in Cancer (https://bioinformatics.mdanderson.org/public-software/tanric/)[21,22] based on TCGA RNA sequencing data on 11 tumor data sets, with estimates of expression levels represented in reads per kilobase million. The significance of the differential expression of each lncRNA, comparing its expression estimates in AA and EA cancer patients, was computed using *U* tests while controlling for the FDR for each tumor type. For invasive breast (BRCA) and uterine corpus endometrial (UCEC) carcinomas, the analyses were performed independently for TNBC and non-TNBC cases and for endometrioid endometrial adenocarcinoma (EEA) and non-EEA cases, respectively. To avoid low signal-to-noise instances, only lncRNAs with mean absolute deviation scores greater than 0·1 were included in the final analysis.

To identify common biologic processes that potentially explain the association between AA race and overall survival, pathway analyses were performed using Ingenuity Pathway Analysis (IPA) software (version 46901286; Ingenuity Systems). After comparing DNA methylation profiles of EA and AA patients for each TCGA tumor type, probes with the most extreme P values were extracted and mapped to human genes. Probes with beta differences greater than 0·1 at an FDR of 0·1 were considered to represent significant changes in methylation. For a gene to which multiple probes were mapped, the probe with the lowest P value was selected to represent that gene. After comparing miRNA expression profiles for each tumor type, only miRNAs significant at an FDR of 0·1 and P<0·05 were mapped. Additional pathway analysis of differentially methylated genes was performed using gene set enrichment analysis, GSEA software, and Molecular Signature Database (Broad Institute and UC San Diego).[23] Depicted pathways have FDR q value < 0·05 and absolute NES > 1·5. Unsupervised hierarchical clustering, heatmap rendering and figure creation was performed in the R programming language (R Foundation for Statistical Computing, Vienna, Austria) using ggplot2 and heatmap packages.

Results

Survival Analysis

We first focused on overall survival differences between AA and EA patients using the SEER data. After applying the exclusion criteria, 2,045,839 patients with complete data (12·5% AA and 87·5% EA) were available for analysis. We sorted them into 42 different cancer type populations, ranging from 361,847 breast carcinoma cases to 188 placental cancer cases. Table S2 in the Supplementary Appendix displays the clinical characteristics of the SEER study patients. Among all patients, AA race was associated with increased risk of death for a majority of the cancer types (Fig. S2 in the Supplementary Appendix). Specifically, 28 cancer types were associated with an increased risk of death, with HRs greater than 1·0. To determine whether the survival outcome is related to sociodemographic factors, we performed a secondary survival analysis using the NCDB data. We focused on 19 cancer types associated with a survival disadvantage and with TCGA data available for molecular analyses. The NCDB data were available on a total of 5,150,023 patients (11·6% AA and 88·4% EA). Table S3 in the Supplementary Appendix displays the clinical characteristics of the NCDB study patients. In the NCDB and SEER survival analyses, we observed persistently higher mortality for 13 tumor types in AA patients (HR>1·0) compared to 6 tumor types in which the higher mortality rate was lowered after adjustment for socioeconomic variables (Fig. 1). We then performed molecular analyses for cancer types with increased mortality seen in both the NCDB and SEER data sets using TCGA data.

MiRNA Analysis

We analyzed genetic ancestry data for 5283 tumors across 13 tumor types, 12·8% of which were obtained from AAs and 87·2% were obtained from EAs. We first compared miRNA expression levels in AA and EA patients. Of the 13 tumor types analyzed, we focused on the 5 types with the highest numbers of significantly differentially expressed miRNAs at an FDR of 0·1: BRCA (300 miRNAs), kidney renal clear cell carcinoma (KIRC; 408 miRNAs), prostate adenocarcinoma (PRAD; 102 miRNAs), thyroid carcinoma (THCA; 63 miRNAs), and UCEC (177 miRNAs). All differentially expressed

miRNAs are listed according to tumor type in Table S4 in the Supplementary Appendix. Unsupervised clustering analysis of the most variable miRNAs in AAs revealed similarities between KIRC, PRAD, THCA and BRCA tumors (Fig. 2a) We performed pathway analysis to determine the clinical significance and biologic functions of the differentially expressed miRNAs for each cancer type. For all cancer types, differentially expressed miRNAs were associated with cancer drug resistance pathways through cancer drug efflux (Fig. S3.). Also for all cancer types, 11 miRNAs were commonly dysregulated in AAs (Table 1). Investigation of these miRNAs within published data indicates their role in oncogenesis. Notably, expression of miRNAs miR-15a, miR-17, and miR-130-3p was frequently elevated in patients with breast, kidney renal clear cell, prostate and thyroid carcinomas.

LncRNA Analysis

To characterize dysregulation of lncRNA expression that may be associated with AA race, we analyzed their expression in patients with 11 cancer types having adequate sample sizes. The dysregulated lncRNAs by cancer type are listed in Table S5 in the Supplementary Appendix. Of these 11 cancer types, only four tumor types had 20 or more dysregulated lncRNAs associated with AA race: THCA (20 lncRNAs), PRAD (80 lncRNAs), BRCA (77 lncRNAs), and KIRC (34 lncRNAs) (Fig. S4 in the Supplementary Appendix). More than 70% of the lncRNAS were unique to specific cancer types.

DNA Methylation

We then analyzed DNA methylation and gene expression data for all cancer types. We identified nine cancer types associated with dysregulated gene methylation in AA race: BLCA (30 genes), BRCA (325 genes), COAD (294 genes), ESCA (541 genes), HNCA (204 genes), KIRC (263 genes), PRAD (269 genes), THCA (482 genes), and UCEC (157 genes). (Table S6 in the Supplementary Appendix). Unsupervised cluster analysis of DNA methylation and gene expression was performed on cancer types with greater than 100 genes affected (Fig 2b). Among seven of eight cancer types, we identified hypomethylation of TRPC5 (transient receptor potential channel C5), S100A14 (S100 calcium binding protein A14) and MIR662 in AA compared to EA tumors leading to

increased gene expression. These alterations are known to elicit resistance to chemotherapy [24,25-27] and constitute potential therapeutic targets. Additionally, we observed that ESCA tumors presented distinct sets of DNA methylation compared to other cancer types with 86% (466 of 541) unique genes affected. This may represent the already known effect of environmental exposure on molecular alterations of esophageal carcinogenesis.[28] Finally, the gene set enrichment analysis of biological pathways enriched in genes with altered DNA methylation in AAs across seven cancer types revealed an enrichment in metabolism, development and signaling pathways (Fig. S5).

Tumor Subtyping

Finally, we compared histologic subtypes in BRCA and UCEC AA patients. In breast carcinoma, TNBC is an aggressive tumor phenotype more frequently diagnosed in AA than in EA women.[29] Similarly, AA women are more likely than EA women to be diagnosed with type 2 (non-EEA) uterine tumors, which are typically more aggressive than other uterine tumor subtypes.[30] Therefore, we compared epigenetic modifications in TNBC versus non-TNBC cases and EEA versus non-EEA cases in AA and EA patients. AA TNBC and non-TNBC cases were associated with 274 and 41 differentially expressed miRNAs, respectively. We performed pathway analysis of individual data sets (TNBC and non-TNBC) and then compared predicted activation of functional pathways using IPA software. AA miRNA expression data in TNBC were associated with an increase in *Angiogenesis* pathway, while the miRNA expression data of AA in non-TNBC tumor types were associated with a decrease in *Angiogenesis* pathway (Table 2). Additionally, all UCECs with dysregulated miRNA expression were of the EEA subtype; therefore, we were not able to compare miRNA expression profiles in the EEA versus non-EEA cases. Expression of TRPC5, the most significantly hypomethylated gene in the BRCA and UCEC data sets, was overexpressed in all histological subtypes. We found significant differences in lncRNA expression associated with AA race only in the non-TNBC and EEA groups (Table S5 in the Supplementary Appendix).

Discussion

Our findings demonstrate that AAs have a higher risk of death for many cancer types compared to EAs. This greater risk of death was consistent across two large cancer patient databases and persisted despite controlling for socioeconomic factors, access to care, and insurance status. Furthermore, we observed that among these cancer types with a persistent survival disadvantage in AAs, epigenetic modifications unique to AAs may have influenced tumor biology and response to therapy.

This knowledge adds to a growing body of evidence that racial disparities may be due to genetic and biological differences.[31] For instance, differential expression of inflammatory mediators such as IL-6, and inflammatory cytokines have been found to be disproportionally increased in AA breast cancer patients.[32] Additionally, mediators of angiogenesis (vascular endothelial growth factor [VEGF]), and immune cells with a tumor-promoting phenotype (tumor-associated macrophages [TAMS]) have both been found to be increased in tumors from AA patients. [33] The genetic basis for these tumor promoting events has yet to be elucidated but suggests inherent racial differences.

The strengths of this study are the use of two large, diverse, nationally representative samples of patients paired with an equally diverse molecular database. We adjusted for several measures of socioeconomic status that could confound the relationship between race and survival. Despite controlling for these variables, we found a persistent cancer survival disadvantage in AA patients. Furthermore, using TCGAA, we were able to reliably assign tumor types to specific races with certainty based on integrated computational algorithms to infer the genetic ancestry of TCGA patients at global and local levels. Previous studies using TCGA data relied on self-identified race or ethnicity, which were challenging and limited the ability to elucidate the genetic contribution to cancer disparities.[34] Additionally, despite emerging interests in the non-coding genome for diagnostic and therapeutic purposes, there remains a need to functionally annotate cancer-associated miRNAs and lncRNAs. Our findings help to identify the molecular roles of epigenetic processes that may contribute to cancer pathology.

Limitations of the present study include the low absolute numbers of patients having some cancer types (lymphoma, melanoma, uterine sarcoma, and glioblastoma multiforme), which restricted our ability to perform molecular analysis for these tumors. In addition, the socioeconomic status of patients in the TCGA data is unavailable, so we were unable to adjust for this variable when comparing racial differences according to tumor biology. Also, the molecular data in TCGA lacked detailed clinical information sufficient to provide adjusted outcome measures. We found higher mortality rates associated with AA race in several cancer types consistent across two large cancer registries, SEER and NCDB. However, we are not able to validate data from either cohort, which is an inherent limitation of these data sources. Moreover, both data sets lack or have only incomplete treatment data such as chemotherapy, surgical intervention and radiation therapy. Without these important variables, we are unable to control for the biological effects of treatment on the epigenome. Finally, while we include proxies for socioeconomic status in our population-based analysis, we are unable to assess systematic differences between races. For example, patients treated in the public hospital setting are found to have a longer interval from diagnosis to surgery and fewer preoperative visits.[35] We also did not analyze trends in survival among other minority groups, including Asian, Native American, and Hispanic cancer patients, some of whom may also have survival disadvantages.[36]

Hispanic patients make up a fast-growing minority group in the US, however the heterogeneity within the group makes molecular analysis particularly challenging. This highlights the need for additional efforts in identifying larger samples of underrepresented patients with comprehensive clinical and genomic profiles to better understand disparities in cancer survival across other minority groups.

In summary, we identified epigenetic modifications unique to AA cancer patients that may have clinical importance due to their effects on tumor aggressiveness, response to therapy, and overall survival.

References:

1. Allard JE, Maxwell GL. Race disparities between black and white women in the incidence, treatment, and prognosis of endometrial cancer. *Cancer Control.* 2009;16(1):53-56.
2. Bach PB, Schrag D, Brawley OW, Galaznik A, Yakren S, Begg CB. Survival of blacks and whites after a cancer diagnosis. *Jama.* 2002;287(16):2106-2113.
3. Howard J, Hankey BF, Greenberg RS, et al. A collaborative study of differences in the survival rates of black patients and white patients with cancer. *Cancer.* 1992;69(9):2349-2360.
4. DeSantis CE, Miller KD, Goding Sauer A, Jemal A, Siegel RL. Cancer statistics for African Americans, 2019. *CA: a cancer journal for clinicians.* 2019.
5. Newman LA. Disparities in breast cancer and african ancestry: a global perspective. *Breast J.* 2015;21(2):133-139.
6. Shavers VL, Brown ML. Racial and ethnic disparities in the receipt of cancer treatment. *J Natl Cancer Inst.* 2002;94(5):334-357.
7. Ward E, Jemal A, Cokkinides V, et al. Cancer disparities by race/ethnicity and socioeconomic status. *CA: a cancer journal for clinicians.* 2004;54(2):78-93.
8. Williams DR, Mohammed SA, Shields AE. Understanding and effectively addressing breast cancer in African American women: Unpacking the social context. *Cancer.* 2016;122(14):2138-2149.
9. Newman LA, Stark A, Chitale D, et al. Association Between Benign Breast Disease in African American and White American Women and Subsequent Triple-Negative Breast Cancer. *JAMA Oncol.* 2017;3(8):1102-1106.
10. Lindner R, Sullivan C, Offor O, et al. Molecular phenotypes in triple negative breast cancer from African American patients suggest targets for therapy. *PloS one.* 2013;8(11):e71915.
11. Yuan J, Hu Z, Mahal BA, et al. Integrated Analysis of Genetic Ancestry and Genomic Alterations across Cancers. *Cancer Cell.* 2018;34(4):549-560.e549.

12. Ademuyiwa FO, Tao Y, Luo J, Weilbaecher K, Ma CX. Differences in the mutational landscape of triple-negative breast cancer in African Americans and Caucasians. *Breast Cancer Res Treat.* 2017;161(3):491-499.
13. Huang FW, Mosquera JM, Garofalo A, et al. Exome Sequencing of African-American Prostate Cancer Reveals Loss-of-Function ERF Mutations. *Cancer Discov.* 2017;7(9):973-983.
14. Wang H, Schmit SL, Haiman CA, et al. Novel colon cancer susceptibility variants identified from a genome-wide association study in African Americans. *Int J Cancer.* 2017;140(12):2728-2733.
15. Feinberg AP. The Key Role of Epigenetics in Human Disease Prevention and Mitigation. *N Engl J Med.* 2018;378(14):1323-1334.
16. Teschendorff AE, West J, Beck S. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Hum Mol Genet.* 2013;22(R1):R7-r15.
17. Ahmad A, Azim S, Zubair H, et al. Epigenetic basis of cancer health disparities: Looking beyond genetic differences. *Biochimica et biophysica acta Reviews on cancer.* 2017;1868(1):16-28.
18. Brewer LC, Redmond N, Slusser JP, et al. Stress and Achievement of Cardiovascular Health Metrics: The American Heart Association Life's Simple 7 in Blacks of the Jackson Heart Study. *J Am Heart Assoc.* 2018;7(11).
19. Grossman RL, Heath AP, Ferretti V, et al. Toward a Shared Vision for Cancer Genomic Data. *N Engl J Med.* 2016;375(12):1109-1112.
20. Pounds S, Morris SW. Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values. *Bioinformatics.* 2003;19(10):1236-1242.
21. Li J, Han L, Roebuck P, et al. TANRIC: An interactive open platform to explore the function of lncRNAs in cancer. *Cancer research.* 2015;75(18):3728-3737.
22. Chiu H-S, Somvanshi S, Patel E, et al. Pan-cancer analysis of lncRNA regulation supports their targeting of cancer genes in each tumor context. *Cell reports.* 2018;23:297-312.
23. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545-15550.
24. He DX, Ma X. Transient receptor potential channel C5 in cancer chemoresistance. *Acta Pharmacol Sin.* 2016;37(1):19-24.
25. Ma X, Chen Z, Hua D, et al. Essential role for TrpC5-containing extracellular vesicles in breast cancer with chemotherapeutic resistance. *Proc Natl Acad Sci U S A.* 2014;111(17):6389-6394.

26. Filipska M, Skrzypski M, Czetyrbok K, et al. MiR-192 and miR-662 enhance chemoresistance and invasiveness of squamous cell lung carcinoma. *Lung Cancer.* 2018;118:111-118.

27. Qian J, Ding F, Luo A, Liu Z, Cui Z. Overexpression of S100A14 in human serous ovarian carcinoma. *Oncol Lett.* 2016;11(2):1113-1119.

28. Loomis D, Guyton KZ, Grosse Y, et al. Carcinogenicity of drinking coffee, mate, and very hot beverages. *The Lancet Oncology.* 2016;17(7):877-878.

29. Sturtz LA, Melley J, Mamula K, Shriver CD, Ellsworth RE. Outcome disparities in African American women with triple negative breast cancer: a comparison of epidemiological and molecular factors between African American and Caucasian women with triple negative breast cancer. *BMC Cancer.* 2014;14(1):62.

30. Cote ML, Ruterbusch JJ, Olson SH, Lu K, Ali-Fehmi R. The Growing Burden of Endometrial Cancer: A Major Racial Disparity Affecting Black Women. *Cancer Epidemiology Biomarkers & Prevention.* 2015.

31. Deshmukh SK, Srivastava SK, Tyagi N, et al. Emerging evidence for the role of differential tumor microenvironment in breast cancer racial disparity: a closer look at the surroundings. *Carcinogenesis.* 2017;38(8):757-765.

32. Deshmukh SK, Srivastava SK, Bhardwaj A, et al. Resistin and interleukin-6 exhibit racially-disparate expression in breast cancer patients, display molecular association and promote growth and aggressiveness of tumor cells through STAT3 activation. *Oncotarget.* 2015;6(13):11231-11241.

33. Martin DN, Boersma BJ, Yi M, et al. Differences in the tumor microenvironment between African-American and European-American breast cancer patients. *PloS one.* 2009;4(2):e4531.

34. Spratt DE, Chan T, Waldron L, et al. Racial/Ethnic Disparities in Genomic Sequencing. *JAMA Oncol.* 2016;2(8):1070-1074.

35. Frey MK, Moss HA, Musa F, et al. Preoperative experience for public hospital patients with gynecologic cancer: Do structural barriers widen the gap? *Cancer.* 2016;122(6):859-867.

36. Hall JE, Moonesinghe R, Bouye K, Penman-Aguilar A. Racial/Ethnic Disparities in Mortality: Contributions and Variations by Rurality in the United States, 2012(-)2015. *Int J Environ Res Public Health.* 2019;16(3).

37. Geretto M, Pulliero A, Rosano C, Zhabayeva D, Bersimbaev R, Izzotti A. Resistance to cancer chemotherapeutic drugs is determined by pivotal microRNA regulators. *Am J Cancer Res.* 2017;7(6):1350-1371.

38. Jiang H, Yu WW, Wang LL, Peng Y. miR-130a acts as a potential diagnostic biomarker and promotes gastric cancer migration, invasion and proliferation by targeting RUNX3. *Oncol Rep.* 2015;34(3):1153-1161.

39. Zhang HD, Jiang LH, Sun DW, Li J, Ji ZL. The role of miR-130a in cancer. *Breast Cancer.* 2017;24(4):521-527.

40. Sueta A, Yamamoto Y, Tomiguchi M, Takeshita T, Yamamoto-Ibusuki M, Iwase H. Differential expression of exosomal miRNAs between breast cancer patients with and without recurrence. *Oncotarget.* 2017;8(41):69934-69944.

41. Ma W, Ma CN, Zhou NN, Li XD, Zhang YJ. Up- regulation of miR-328-3p sensitizes non-small cell lung cancer to radiotherapy. *Sci Rep.* 2016;6:31651.

42. Feng Y, Bai F, You Y, et al. Dysregulated microRNA expression profiles in gastric cancer cells with high peritoneal metastatic potential. *Exp Ther Med.* 2018;16(6):4602-4608.

43. Shi Q, Zhou Z, Ye N, Chen Q, Zheng X, Fang M. MiR-181a inhibits non-small cell lung cancer cell proliferation by targeting CDK1. *Cancer Biomark.* 2017;20(4):539-546.

44. Meijer LL, Garajova I, Caparello C, et al. Plasma miR-181a-5p Downregulation Predicts Response and Improved Survival After FOLFIRINOX in Pancreatic Ductal Adenocarcinoma. *Ann Surg.* 2018.

45. Pignot G, Cizeron-Clairac G, Vacher S, et al. microRNA expression profile in a large series of bladder tumors: identification of a 3-miRNA signature associated with aggressiveness of muscle-invasive bladder cancer. *Int J Cancer.* 2013;132(11):2479-2491.

46. Lin C, Li Z, Chen P, et al. Oncogene miR-154-5p regulates cellular function and acts as a molecular marker with poor prognosis in renal cell carcinoma. *Life Sci.* 2018;209:481-489.

47. Aqeilan RI, Calin GA, Croce CM. miR-15a and miR-16-1 in cancer: discovery, function and future perspectives. *Cell Death And Differentiation.* 2009;17:215.

48. Zhu Y, Gu J, Li Y, et al. MiR-17-5p enhances pancreatic cancer proliferation by altering cell cycle profiles via disruption of RBL2/E2F4-repressing complexes. *Cancer Lett.* 2018;412:59-68.

49. Egawa H, Jingushi K, Hirono T, et al. The miR-130 family promotes cell migration and invasion in bladder cancer through FAK and Akt phosphorylation by regulating PTEN. *Sci Rep.* 2016;6:20574.

50. Duan J, Zhang H, Qu Y, et al. Onco-miR-130 promotes cell proliferation and migration by targeting TGFbetaR2 in gastric cancer. *Oncotarget.* 2016;7(28):44522-44533.

Tables

Table 1. MiRNAs Commonly Dysregulated in AA Cancer Patients

MiRNA Component	Expression level in AAs					Characteristics
	UCEC	THCA	PRAD	BRCA	KIRC	
miR-27a	Low	---	---	High	High	Resistance to 5-Fluorouracil mediated by miR-27a/b[37]
miR-331-5p	Low	---	---	---	---	---
miR-130a	High	High	---	High	High	Potential diagnostic marker for breast, gastric, ovarian carcinomas [38-40]
miR-328	Low	---	---	---	---	Overexpression can improve radiosensitivity of NSCLC[41]
miR-181a	---	High	High	High	High	Regulatory role in NSCLC pancreatic, gastric, and colon carcinomas[42-44]
miR-133a	---	---	---	---	High	Associated with aggressive bladder carcinoma[45]
miR-379	---	---	---	High	High	---
miR-154	Low	---	High	Low	High	Prognostic significance for renal cell carcinoma[46]
miR-15a	High	High	High	High	High	miRNAs encoded by miR-15/16 may function as tumor suppressors[47]
miR-17	High	High	High	High	High	Overexpressed in various solid tumors[48]
miR-103-	High	High	High	High	High	Considered an onco-miR in

3p	gastric, and bladder carcinomas[49,50]

NSCLC, non-small cell lung cancer.

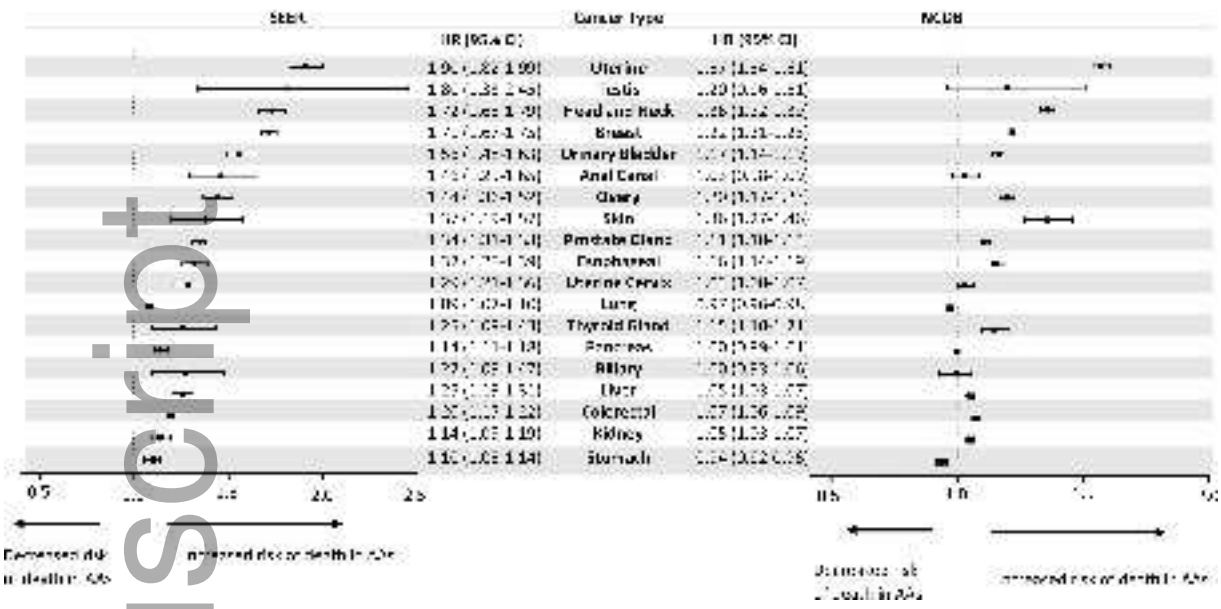
Table 2: Disease and Functional Networks in AA Patients with TNBC and Non-TNBC Based on miRNA Expression.

Tumor Type	Diseases or functions annotation	P-value	Predicted Activation	Z-Score	Associated microRNAs
TNBC					
	Angiogenesis	1.64 x 10^{-03}	Increased	0·748	mir-21, mir-27, mir-29, mir-361
Non-TNBC					
	Angiogenesis	2.99 x 10^{-09}	Decreased	-1·890	mir-135, mir-17, mir-154, mir-320, mir-19, mir-130, mir-181, mir-185, mir-1301, mir-217, mir-181a-5p, mir-15, mir-221, mir-361, mir-8, mir-103, mir-132, mir-146, let-7, mir-25, mir-30, mir-21, mir-320b

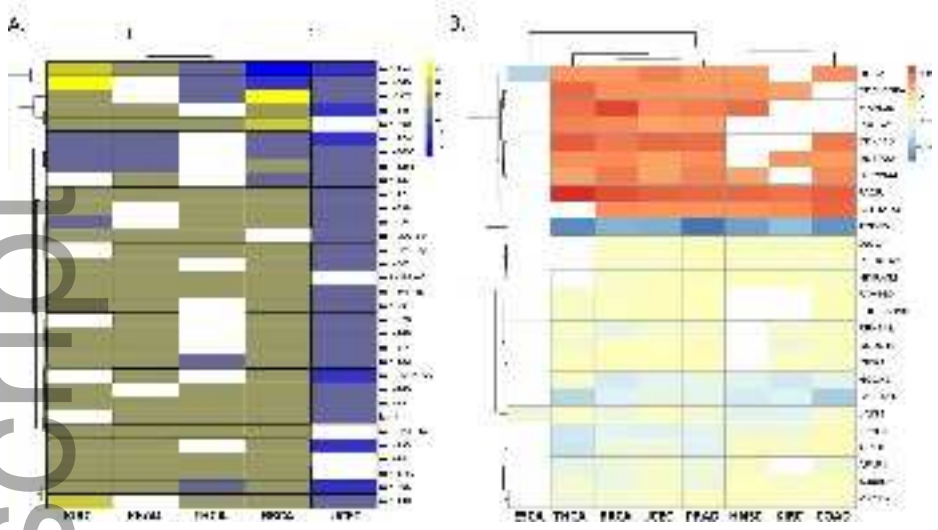
Figure Legends:

Figure 1. Risk of Death in AA Cancer Patients According to Tumor Type.
Shown are results of multivariate analysis of the risk of death in AA cancer patients in the SEER and NCDB databases. CI, confidence interval.

Figure 2. TCGA features of AA patients. Unsupervised hierarchical clustering and heatmap of top commonly dysregulated microRNA (A) and DNA methylation targets (B) in AA versus EA cancer patients from 5283 tumors. (A) Fold change in AA vs EA patients for significant common microRNAs in 5 cancer types. (B) Delta beta value in AA vs EA patients for top loci. Red boxes indicate relative hyper methylation in AA patients. Blue boxes note relative hypo methylation in AA patients. White boxes represent genes with no significant alteration in AA versus EA patients.



cncr_32598_f1.tif



cncr_32598_f2.tif