

Epigenetic analysis identifies factors driving racial disparity in prostate cancer

Richa Rai^{*1}, Shalini S Yadav^{*1}, Heng Pan², James O'Connor¹, Mohamed Alshalalfa³, Elai Davicioni³, Emanuela Taioli⁴, Olivier Elemento², Ashutosh K Tewari^{#1}, Kamlesh K Yadav^{#1}

¹ Department of Urology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

² Department of Physiology and Biophysics, Institute for Precision Medicine, Weill Cornell Medical College, New York, NY, USA

³ GenomeDx Biosciences, Vancouver, British Columbia, Canada

⁴ Department of Population Health Science and Policy and Institute for Translational Epidemiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

* Contributed equally

Corresponding authors: ash.tewari@mountsinai.org, kamlesh.yadav@mountsinai.org.

Abstract

Prostate cancer (PCa) is the second most leading cause of death in men worldwide. African American men (AA) represent more aggressive form of PCa as compared to Caucasian (CA) counterparts. Evidence suggests that genetic and other biological factors could account for the observed racial disparity. We analyzed the cancer genome atlas (TCGA) dataset (2015) for existing epigenetic variation in AA and CA prostate cancer patients, and carried out Reduced Representation Bisulphite Sequencing (RRBS) analysis to identify global methylation changes in AA and CA prostate cancer patients. The TCGA dataset analysis revealed that the epigenetic heterogeneity could be categorized into 4 classes, where AA associated primarily to methylation cluster 1 (p value 0.048), and CA associated to methylation cluster 3 (p value 0.000146). We identified enrichment of Wnt signaling genes in both AA and CA, however they were differentially activated in terms of canonical and non-canonical Wnt signaling pathway activation. This was further validated using the GenomeDx expression data. Our RRBS data also suggested distinct methylation patterns in AA compared to CA, and in part validated our TCGA findings. Survival analysis using the RRBS data suggested hypomethylated genes to be significantly associated with recurrence of prostate cancer in CA ($p=6.07 \times 10^{-6}$) as well as in AA ($p=0.0077$). Overall, the observed racial disparity in the molecular mechanism involved in the pathogenesis of prostate cancer suggests diverse heterogeneity that potentially could affect survival and should be considered during prognosis and treatment.

Keywords: Prostate Cancer, African-American, Caucasian, Epigenetics

Introduction

Prostate cancer (PCa) is one of the most prevalent malignancies affecting men (Fitzmaurice et al., 2013) [1] and the second-most leading cause of cancer-related death for men in the US [2,3]. According to Surveillance, Epidemiology, and End Results (SEER) there is a decline in prostate cancer incidence over the past few decades [4]. However, African American men (AA) present early onset of the disease with high-grade tumor, aggressive metastasis potential with increased risk of recurrence, and have worse survival outcomes when compared to their CA counterparts [5-8]. It is suggested that racial disparity in PCa is influenced by socioeconomic status, diet and lifestyle [9-11], but several lines of evidence indicate that biological factors like genetic and epigenetic factors are major contributors to this disparity [8, 12]. Such factors include ERG rearrangement, PTEN deletion, SPINK1 overexpression, SPOP mutation and 8q24 SNPs [13, 14]. Furthermore, PCa in AA is associated with higher testosterone levels and prominent immune related tumor biology [15-17], with tumor microenvironment dynamics associated with more aggressive phenotype in AA when compared to CA [18].

Aberrant DNA methylation is an important epigenetic modulation that has been implicated in PCa etiology and disease progression [19]. Methylation changes primarily occur in CpG islands sequestered in the promoter regions [20]. Whole genome methylation studies have identified a large number of hypo and hypermethylated regions in PCa patients [21, 22]. However, previous methylation-based racial disparity studies are limited and have either focused only on a specific gene or benign tissues [23, 24]. More recently, a methylation study using arrays reported differential methylation changes in ABCG5, MST1R, SNRPN and SHANK2 genes in AA and CA PCa patients [25]. They observed ABCG5 and SNRPN genes to be associated with the proliferation and invasion in cell lines of CA PCa patients [25].

In this study, we investigated the epigenetic variations in AA and CA PCa patients and how they are associated with survival and recurrence outcomes using the cancer genome atlas (TCGA) dataset [26] and Reduced Representation Bisulfite Sequencing (RRBS).

Materials and Methods

Prostate Tissue Samples

Prostate tumor tissues and adjacent benign prostate tissues were collected from patients at Icahn School of Medicine at Mount Sinai as per Institutional Review Board approved protocols (GCO# 06-0996, 14-0318 and surgical consent). The fresh frozen prostatectomy specimens were procured after the tumor and benign regions were identified and marked by the pathologist. Three AA and three CA PCa patients were enrolled in this study, their clinical profile including tumor stage and Gleason score has been elaborated in **Table 1**. Genomic DNA was extracted from frozen tissues using Qiagen's DNA/RNA extraction kit (Cat No. 80204) following the manufacturer's instruction.

Reduced Representation Bisulphite Sequencing

Library preparation for RRBS was performed using NuGEN Ovation® RRBS Methyl-Seq System (Cat NO: 0353,0553) for Illumina Sequencing. In brief, the preparation steps included: 1) DNA fragmentation by MspI enzyme digestion; 2) ligation of adapter; 3) end repair; 4) DNA oxidation; 5) Bisulphite conversion; 6) Desulfonation and purification; and 7) Amplification and purification. The libraries were sequenced using Illumina HiSeq and MiSeq platform (75bp paired end) at The Genomics Core Facility at the Icahn Institute and Department of Genetics and Genomic Sciences.

Computational analysis

Bisulphite sequencing data was analyzed using Bismark v0.15.0 software. It is a set of tools for the Bisulfite-Seq data which performs alignments of bisulfite-treated reads to a reference genome as well as cytosine methylation calls at the same time. Bisulphite converted hg19 reference genome was used for the alignment and mapping of reads. Sequence reads that produce a unique alignment against the bisulfite genomes are then compared to the normal genomic sequence and the methylation state of all cytosine positions in the read is inferred. Read alignment details for RRBS data is summarized in **Table 2**. The methylation status of the uniquely mapped reads was extracted using the `bismark_methylation_extractor` script in Bismark. Only CpGs with at least five reads covering them were used for downstream analysis. Methylation of a specific region was calculated by averaging the methylation levels of all CpGs covered in that region. We calculated promoter methylation levels for each sample then compared the variation among CA vs AA and tumor vs benign region for each racial group. Promoters were defined as ± 2 kb windows centered on RefSeq transcription start sites.

Survival analysis

SurvExpress analysis tool (<http://bioinformatica.mty.itesm.mx:8080/Biomatec/SurvivaX.jsp>) was used to determine the association between hypermethylated genes and survival. This tool uses both the gene list from our dataset and publically available human cancer database to predict biomarkers for risk assessment and survival. PRAD TCGA database [26] was used for the identification of biomarkers associated with survival. The Sobner [27] and Tylor databases [28] were used for the analysis of survival and recurrence. The software generates risk groups by either by splitting the ordered Prognostic Index (risk score estimated by beta coefficients multiplied by gene expression values) for each samples at the median where samples are distributed equally in both groups or by using an optimized algorithm from the ordered

prognostic index [29]. The survexpress software was used in default setting and it performs a log-rank test along all values of the arranged PI and then, the algorithm selects the split point where the p-value is minimum. This procedure is repeated until no changes are observed. This tool presents the output in the form of Kaplan-Meier curve for risk groups, concordance index which estimate the probability that subjects with a higher risk will experience the event after subjects with a lower risk and p-value of the log-rank testing equality of survival curves.

TCGA data re-analysis

The Cancer Genome Atlas Network (2015) performed extensive genomic, transcriptomic, epigenetic and proteomics studies on a large cohort of prostate cancer patients (333 primary prostate tumor tissues) to identify the molecular and biological heterogeneity in prostate cancer patients [26]. In order to investigate the epigenetic variation and specific methylation pattern in AA and CA men, we retrieved the TCGA datasets with methylation cluster information and clinical details for each patient from the cBioPortal for Cancer Genomics (http://www.cbioportal.org/study.do?cancer_study_id=prad_tcga_pub). Of the 333 prostate cancer patients who were enrolled in TCGA study, 43 PCa patients were reported to be AA and 162 were CA. Further, to identify the epigenetic changes existing in each racial group we used gene lists that were reported to be dysregulated in each methylation cluster in TCGA study. The list of differentially regulated genes in each cluster of TCGA study was procured from firebrowse (<http://firebrowse.org/?cohort=PRAD>) [30]. We also used TCGA methylated data list to validate our RRBS analysis, which was retrieved from (<http://firebrowse.org/?cohort=PRAD>) [portal](#) [31].

Functional analysis

The Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, Redwood City, CA) was used to identify the relevant bio-functions and the pathways associated with differentially expressed gene in each methylation cluster. The software generates pathways utilizing the list of genes from TCGA data and the list of genes stored in the ingenuity knowledge database, which is based on functional annotations and experimental observations. It also computes a p value for each pathway, to assess the likelihood of the association between the focus genes and a canonical pathway being not random. IPA also predicts the activation or inhibition of a canonical pathway based on z-score, which is automatically calculated based on differentially expressed genes in both the uploaded dataset and the information stored in IPA knowledge database. A positive z-score suggests the activation, whereas a negative z-score indicates inactivation of the pathway.

GenomeDx analysis

Genomics Resource Information Database (GRID) patients' samples were used for the validation of our observation of differential Wnt signaling in CA and AA cohorts from TCGA data. RNA was extracted from routine formalin-fixed, paraffin embedded (FFPE) Radical prostatectomy tumor tissues, which were amplified and hybridized to Human Exon 1.0 ST microarrays (Thermo-Fisher, Carlsbad, CA). All GRID data were normalized utilizing the single channel array normalization algorithm [32]. GRID samples were processed as described in [33].

Results

Identification of differential methylation clusters among AA and CA PCa patients in TCGA dataset

An integrative epigenetic and genetics analysis in TCGA study identified four distinct methylation groups which are named as methylation cluster (MC) 1, 2, 3 and 4 [26]. Prevalent PCa genetic alterations like ERG, ETV1, ETV4 fusions and SPOP, FOXA1, IDH1 mutations

were reported to be differentially associated with each methylation clusters [26]. Interestingly, ERG fusion-positive PCa patients were differentially distributed in MC1 and MC3; one-third of them exhibited hypermethylation and belonged to methylation cluster (MC1) whereas, two-third of ERG positive PCa patients had hypomethylated DNA loci (MC3) **Figure 1A**. MC2 had patients with SPOP, FOXA1 and IDH1 mutations. They also included ETV1 and ETV4 fusion and had highly methylated DNA. MC4 was positive for ERG, ETV1 and ETV4 fusion as well as SPOP and FOXA1 mutations.

In this study, we re-analyzed TCGA data for identification of epigenetic variation among AA and CA racial groups. PCa patients were then divided into two groups based of their race (43 AA and 162 CA) and we observed that AA men were more likely to present the MC1 pattern (p value 0.048), while CA men mostly presented with MC3 (p value 0.000146), **Figure 1B**. The MC1 cluster displays a distinct spectrum of hypermethylated loci whereas the MC3 cluster shows hypomethylated loci.

Pathways associated with different methylation clusters

To study the predominant signaling pathways associated in each methylation cluster we retrieved the mRNA expression data from TCGA and generated a list of differentially expressed gene in each methylation cluster. We used the ingenuity pathway analysis (IPA) to identify pathways affected in each methylation cluster. We observed Wnt signaling pathway to be differentially regulated in all four clusters suggesting it to play an important role in PCa pathogenesis (**Supplementary Table 1**). Since AA and CA PCa patients distributed distinctly in MC1 and MC3 only we focused our analysis on these two clusters only. The Analysis of MC1 showed that non-canonical Wnt signaling (Wnt/Ca⁺² signaling) was downregulated in CA whereas it was activated in AA, **Figure 2A**. The analysis of MC3 showed exclusive activation of non-canonical

Wnt/Ca⁺² in CA. Furthermore, inflammation/immune cell related pathways and PI3K signaling was observed to be associated with upregulated genes in MC1 (whereas down regulated genes associated with cell cycle signaling). In the case of MC3, amino acid metabolism related pathways were associated with upregulated genes, and downregulated genes were associated with cytokine signaling and PI3K-pathway alterations **Table 3**.

The finding of differential regulation of Wnt signaling in MC1 (representation of AA cohort) and MC3 (representation of CA cohort) from TCGA data re-analysis were further validated in a separate and larger cohort of PCa patients among these two racial groups. We used the GRID data from GenomeDx, which uses RNA expression analysis from prostate tumors to provide a Decipher^R genomic score that predicts disease aggressiveness. In this data base, genes belonging to the non-canonical Wnt/Ca⁺² signaling pathway were significantly upregulated in CA as compared to AA (p value 5.6×10^{-06}), **Figure 2B**.

Epigenetic variation in AA and CA Prostate Cancer patients

To further validate our TCGA analysis and identify the epigenetic differences in AA and CA prostate tumors, we performed RRBS analysis on 6 PCa patients, three in each group, **Table 1**. Benign tissue adjacent to tumor was procured from each patient to serve as control. We compared the methylation changes in CA and AA tumor tissues and observed very distinct methylation pattern in both the racial groups. Further, analysis revealed hypermethylation of 33 genes in AA, 19 hypermethylated genes in CA (**Supplementary table 2**). This suggests hypermethylation to be a predominant event in AA, which is in concordance with our observation with TCGA data where MC1 cluster had more hypermethylated genes compared to MC3 **Figure 3A**. Moreover, we observed significant methylation changes in benign vs tumor tissue in both CA and AA **Figure 3B and C**, respectively.

Correlation of RRBS analysis with TCGA data

To validate our finding we matched the methylated gene list from RRBS with the TCGA methylation data. We used the negative correlation between DNA methylation and mRNA expression data from TCGA for this comparison [31]. Methylation mean with cut-off < 0.2 was used for hypomethylated genes and cut-off > 0.3 was used for hypermethylated genes [26]. Even though different platforms were used for the methylation studies, in both the cases, we observed similar methylation patterns; RPS29, ZNF180, POLD4, AKIRIN2, SLC41A3, PDIA3, CPSF2, and SLC35C2 were found to be hypomethylated in AA men in RRBS analysis as well as in TCGA analysis, **Table 4**. Hypermethylated genes from both the analysis include IDH3B, LRP10, BLVRB, ZNF256, LIME1, CD72, URM1, PSAP, PIKFYVE, ARL6IP1 and BTBD2, **Table 5**.

Correlation of hypomethylated genes with survival in AA and CA Prostate cancer patients

To identify whether the genes we identified had any prognostic significance, we performed survival analysis on the hypomethylated genes obtained from RRBS analysis in both the racial groups. PRAD-TCGA (TCGA) dataset (497 samples) [26] and Sboner datasets (281 samples) [27] were used for the prediction of survival outcome whereas Taylor datasets (140 Sample) [28] was used to access the recurrence of prostate cancer in different racial group. Both Sboner data and TCGA data were able to separate the risk group based on differential gene expression listed in **Table 6**. Nonetheless, p value for the risk group separation was better for Sboner dataset as compared to TCGA data for both the racial groups (**Figure 4A, B, Table 6**). In addition, Taylor datasets suggested significant association of hypomethylated genes with biochemical recurrence of prostate cancer in CA (p value 6×10^{-6}) as well as in AA (p value 0.0077) (**Figure 4C, Table 6**).

Discussion

Comprehensive “omics” studies of Prostate cancer in the cancer genome atlas (TCGA) study indicated that diverse heterogeneity exists in PCa patients [26]. Besides, this fact there exist disproportionate incidence, progression and mortality of PCa patients in relation to race. In this study we identified that AA and CA have distinct methylation patterns. Further, pathway analysis revealed that the specific methylation events in each group are associated with different biological functions. Each racial group is associated with differential Wnt signaling and other different biological pathways, which have not been explored in details previously with respect to racial disparity. We also observed discrete methylation changes of genes in AA and CA patients which may contribute to differences in prostate cancer outcomes between these two populations. Epigenetic changes are an inherent genomic property that appears early during development [34] or sometime acquired during lifetime due to exposure of other environmental factors [35, 36]. Thus it is a critical factor in the pathogenesis of PCa and could contribute towards aggressive clinical outcomes.

Several studies reported that hypermethylation of CpG is the predominant event occurring in case of prostate cancer [26, 37]. There are large variations in methylation of specific genes among AA and CA ethnic groups [22-24], reports suggest the existence of race specific methylation differences at the birth [36, 38]. Using the TCGA data and RRBS analysis, we observed that AA have predominant methylation events as compared to CA, and both the groups have very distinct spectrum of methylated loci. This result of race specific global methylation difference is interesting and requires further investigation to elucidate the mechanism driving this disparity.

Further, pathway analysis also suggests disparity in terms of biological functions and signaling mechanism which is a crucial event. For instance, we found canonical and non-canonical Wnt signaling pathways are differentially regulated in AA (MC1) and CA (MC3). Wnt/ β -catenin signaling is upregulated and Wnt/ Ca^{+2} downregulated in AA whereas; the later is downregulated in CA. Although, multiple studies have identified association of Wnt signaling with aggressive, late stage disease, and metastasis in prostate cancer [39-41], to our knowledge, the distinction of Wnt/ β -catenin and Wnt/ Ca^{+2} signaling pathway association with ethnicity in prostate cancer has not been reported previously. Of the two Wnt signaling pathways, Wnt/ β -catenin is very well studied pathway and plays important role in progression of prostate cancer towards metastasis, and development of disease [42, 43]). More recently, non-canonical Wnt (Wnt/ Ca^{+2}) signaling has also been implicated in the pathogenesis of prostate cancer [41, 44]. The involvement of Wnt/ β -catenin pathway in progression could be attributed to its function in proliferation and survival of cancer stem cells and was suggested to act as molecular markers for predicting resistance to several androgen dependent therapies [45]. Lately, genes involved in canonical and non-canonical Wnt signaling pathways are emerging as promising targets for therapeutic intervention in prostate cancer [46, 47]. It would be interesting to explore its therapeutic role in prostate cancer patients of different ethnicities.

We also observed that inflammation related pathway, PI3K signaling, cell cycle signaling, and amino acid metabolism pathways vary differentially in these racial groups. Previous studies have suggested association of interleukins signaling and inflammation to be more prevalent in AA relative to CA men [16, 48]. Also, PI3K signaling is crucial in PCa as Phosphatase and tensin homologue (PTEN) loss, resulting in constitutive activation of the PI3K pathway, is common in advanced prostate cancer. Using reciprocal-pairing of miRNA-mRNA, Wang et al., 2015

predicted the overt activation of PI3K/AKT pathway in African American men compared to CA men [49]. Our observations are similar to these findings. Furthermore, we found cell cycle signaling was downregulated in AAs whereas a recent study shows enhanced cell cycle related transcription in specific subpopulation of CA derived prostate cancer cell-line [50]. Finally, similar to our observation of amino-acid metabolism to be associated with CA, Putluri et al., 2011 reported higher amino acid metabolism in CA derived prostate cancer cell lines treated with androgens [51].

From the present study we also identified distinct set of methylated genes in AA and CA men using RRBS. We observed genes hypermethylated in AA were hypomethylated in CA and vice-versa. Some genes were also found to follow the same methylation pattern as observed in TCGA study which validates our finding. These include the PDIA3 gene which was hypomethylated in AA (hypermethylated in Ca) as well as in TCGA data. Evidence suggests its association with malignant stage of prostate cancer which could be used as potential biomarkers [52]. Furthermore, hypermethylated genes in AA include isocitrate dehydrogenase beta-subunit (IDH3B) and Phosphoinositide Kinase, FYVE-Type Zinc Finger Containing (PIKFYVE), which a study suggests could serve as a urine-based biomarker for prostate cancer [53]. The prostate has been suggested to synthesize high amount of citrate and mutations in isocitrate dehydrogenases have been associated with PCa [54]. The other gene PIKFYVE induces exosome secretion and is known to influence cancer related pathways like metastasis and drug resistance development. [55]. However, the importance of these genes with respect to racial disparity in PCa needs to be explored further.

Using in-silico analysis we observed that the Sboner dataset [27] and Tylor datasets [28] predicted survival and recurrence outcome in both the ethnic group with the RRBS data.

However, lack of ethnic information in the available dataset limits the predicting impact of these genes in specific ethnic populations. Thus, study involving large cohort of AA and CA Prostate cancer patient is needed for the better prediction of bio- markers for risk assessment, aggressive cancer prediction recurrence and survival outcome which will eventually facilitate clinical translation of genome-based prognostic biomarkers.

This is an era of personalized treatment and genomic, epigenomic and proteomic approaches has contributed a lot in the stratification of various complexities with respect to active surveillance, surgery, therapeutics in prostate cancer. The present study indicates that prostate cancer patient of different ethnicity have distinct epigenetic alterations and biological functions, and thus should be managed differently. Further, study with larger cohort of AA and CA PCa is needed with detailed clinical information to predict better biomarkers for the outcomes and for the targeted therapeutic interventions.

Acknowledgements

We are thankful to the Department of Hematology and Oncology, Icahn School of Medicine at Mount Sinai, New York for use of their equipment and resources. Both SSY and KKY thank the Prostate Cancer Foundation for Young Investigator Awards.

Conflict of interest

None.

Grant support

The research work was supported by funds from the Prostate Cancer Foundation and Deane Prostate Health, Icahn School of Medicine at Mount Sinai, NY

Reference

- [1] Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, et al. The Global Burden of Cancer 2013. *JAMA Oncol.* 2015;1:505-527. doi: 10.1001/jamaoncol.2015.0735. PMID: 26181261
PMCID: PMC450082.
- [2] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016; 66:7-30. doi: 10.3322/caac.21332. PMID: 26742998.
- [3] American cancer society (2017). <https://cancerstatisticscenter.cancer.org/#/>
- [4] Noone AM, Howlander N, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). *SEER Cancer Statistics Review, 1975-2015*, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2015/, based on November 2017 SEER data submission, posted to the SEER web site, April 2018.
- [5] Powell IJ, Banerjee M, Sakr W, Grignon D, Wood DP Jr, Novallo M, Pontes E. Should African-American men be tested for prostate carcinoma at an earlier age than white men? *Cancer.* 1999;85:472–477. PMID: 10023717.
- [6] Powell IJ, Banerjee M, Novallo M, Sakr W, Grignon D, Wood DP, Pontes JE. Prostate cancer biochemical recurrence stage for stage is more frequent among African American than white men with locally advanced but not organ-confined disease. *Urology.* 2000;55:246–251. PMID: 10688088.
- [7] Powell IJ, Bock CH, Ruterbusch JJ, Sakr W. Evidence supports a faster growth rate and/or earlier transformation to clinically significant prostate cancer in black than in white American men, and influences racial progression and mortality disparity. *J Urol.*

2010;183:1792-1796. doi:10.1016/j.juro.2010.01.015. PMID: 20299055 PMCID: PMC3840791.

[8] Powell IJ, Bollig-Fischer A. Minireview: the molecular and genomic basis for prostate cancer health disparities. *Mol Endocrinol.* 2013;27:879–91. PMID: 23608645 PMCID: PMC5415277.

[9] Reddy S, Shapiro M, Morton R Jr, Brawley OW. Prostate cancer in black and white Americans. *Cancer Metastasis Rev.* 2003;22:83-86. PMID: 12716039.

[10] Du XL, Fang S, Coker AL, Sanderson M, Aragaki C, Cormier JN, Xing Y, et al. Racial disparity and socioeconomic status in association with survival in older men with local/regional stage prostate carcinoma: findings from a large community-based cohort. *Cancer.* 2006;106:1276-1285. PMID: 16475208.

[11] Bhardwaj A, Srivastava SK, Khan MA, Prajapati VK, Singh S, Carter JE, Singh AP. Racial disparities in prostate cancer: a molecular perspective. *Front Biosci (Landmark Ed).* 2017;22:772-782. PMID: 27814645 PMCID: PMC5242333

[12] Castro P, Creighton CJ, Ozen M, Berel D, Mims MP, Ittmann M. Genomic profiling of prostate cancers from African American men. *Neoplasia.* 2009;11:305–312. PMID: 19242612 PMCID: PMC2647733.

[13] Troutman SM, Sissung TM, Cropp CD, Venzon DJ, Spencer SD, Adesunloye BA, et al. Racial disparities in the association between variants on 8q24 and prostate cancer: a systematic review and meta-analysis. *Oncologist.* 2012;17:312-320. doi: 10.1634/theoncologist.2011-0315. PMID: 22382457 PMCID: PMC3316915.

[14] Khani F, Mosquera JM, Park K, Blattner M, O'Reilly C, MacDonald TY, Chen Z, et al. Evidence for Molecular Differences in Prostate Cancer between African American and

- Caucasian Men. *Clin Cancer Res.* 2014;20:4925-4934. doi: 10.1158/1078-0432.CCR-13-2265. PMID: 25056375 PMCID: PMC4167562.
- [15] Ross R, Bernstein L, Judd H, Hanisch R, Pike M, Henderson B. Serum testosterone levels in healthy young black and white men. *J Natl Cancer Inst.* 1986; 76:45-48. PMID: 3455741.
- [16] Wallace TA, Prueitt RL, Yi M, Howe TM, Gillespie JW, Yfantis HG, Stephens RM, et al. Tumor immunobiological differences in prostate cancer between African-American and European-American men. *Cancer Res.* 2008;68:927-936. doi: 10.1158/0008-5472.CAN-07-2608. PMID: 18245496.
- [17] Winters SJ, Brufsky A, Weissfeld J, Trump DL, Dyky MA, Hadeed V. Testosterone, sex hormone-binding globulin, and body composition in young adult African American and Caucasian men. *Metabolism.* 2001;50:1242-1247. PMID: 11586501.
- [18] Vijayakumar S, Henegan JC, Zhang X, Wang W, Day WA, Vijayakumar V, Moreno CS, Gomez CR. Enriching gene expression profiles will help personalize prostate cancer management for African-Americans: A perspective. *Urol Oncol.* 2017;35:315-321. doi: 10.1016/j.urolonc.2017.04.003. PMID: 28476530.
- [19] Nelson WG, De Marzo AM, and Yegnasubramanian S (2009). Epigenetic alterations in human prostate cancers. *Endocrinology.* 2009;150:3991–4002. doi: 10.1210/en.2009-0573. PMID: 19520778 PMCID: PMC2736081.
- [20] Bernstein BE, Meissner A, and Lander ES. The mammalian epigenome. *Cell.* 2007;128:669–681. PMID: 17320505.
- [21] Kim JH, Dhanasekaran SM, Prensner JR, Cao X, Robinson D, Kalyana-Sundaram S, Huang C, et al. Deep sequencing reveals distinct patterns of DNA methylation in

- prostate cancer. *Genome Res.* 2011;21:1028–1041. doi: 10.1101/gr.119347.110. PMID: 21724842 PMCID: PMC3129246.
- [22] Mahapatra S, Klee EW, Young CY, Sun Z, Jimenez RE, Klee GG, Tindall DJ, et al. Global methylation profiling for risk prediction of prostate cancer. *Clin Cancer Res.* 2012;18:2882–2895. doi: 10.1158/1078-0432.CCR-11-2090. PMID: 22589488.
- [23] Enokida H, Shiina H, Urakami S, Igawa M, Ogishima T, Pookot D, et al. Ethnic group-related differences in CpG hypermethylation of the GSTP1 gene promoter among African-American, Caucasian and Asian patients with prostate cancer. *Int J Cancer.* 2005;116:174-181. PMID: 15800905.
- [24] Kwabi-Addo B, Wang S, Chung W, Jelinek J, Patierno SR, Wang BD, et al. Identification of differentially methylated genes in normal prostate tissues from African American and Caucasian men. *Clin Cancer Res.* 2010;16:3539-3547. doi: 10.1158/1078-0432.CCR-09-3342. doi: 10.1158/1078-0432.CCR-09-3342. PMID: 20606036.
- [25] Devaney JM, Wang S, Furbert-Harris P, Apprey V, Ittmann M, Wang BD, et al. Genome-wide differentially methylated genes in prostate cancer tissues from African-American and Caucasian men. *Epigenetics.* 2015;10:319–328. doi: 10.1080/15592294.2015.1022019. PMID: 25864488 PMCID: PMC4622564.
- [26] Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. *Cell.* 2015;163:1011-1025. doi: 10.1016/j.cell.2015.10.025. PMID: 26544944 PMCID: PMC4695400.
- [27] Sboner A, Demichelis F, Calza S, Pawitan Y, Setlur SR, Hoshida Y, et al. Molecular sampling of prostate cancer: a dilemma for predicting disease progression. *BMC Med*

- Genomics. 2010;3:8. doi: 10.1186/1755-8794-3-8. PMID: 20233430 PMCID: PMC2855514.
- [28] Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010;18:11-22. doi: 10.1016/j.ccr.2010.05.026. PMID: 20579941 PMCID: PMC3198787.
- [29] Aguirre-Gamboa R, Gomez-Rueda H, Martínez-Ledesma E, Martínez-Torteya A, Chacolla-Huaringa R, Rodriguez-Barrientos A, et al. SurvExpress: an online biomarker validation tool and database for cancer gene expression data using survival analysis. *PLoS One*. 2013;8:e74250. doi: 10.1371/journal.pone.0074250. eCollection 2013. PMID: 24066126 PMCID: PMC3774754.
- [30] Broad Institute TCGA Genome Data Analysis Center (2016): Clustering of Methylation: consensus NMF. Broad Institute of MIT and Harvard. doi:10.7908/C16H4GW2.
- [31] Broad Institute TCGA Genome Data Analysis Center (2016): Correlation between mRNA expression and DNA methylation. Broad Institute of MIT and Harvard. doi:10.7908/C1QF8SB3.
- [32] Piccolo SR, Sun Y, Campbell JD, Lenburg ME, Bild AH, Johnson WE. A single-sample microarray normalization method to facilitate personalized-medicine workflows. *Genomics*. 2012;100:337-44. doi: 10.1016/j.ygeno.2012.08.003. PMID: 22959562 PMCID: PMC3508193.
- [33] Spratt DE, Zhang J, Santiago-Jiménez M, Dess RT, Davis JW, Den RB, et al. Development and Validation of a Novel Integrated Clinical-Genomic Risk Group

Classification for Localized Prostate Cancer. *J Clin Oncol.* 2018;36:581-590. doi: 10.1200/JCO.2017.74.2940. PMID: 29185869.

- [34] Nelson WG, Yegnasubramanian S, Agoston AT, Bastian PJ, Lee BH, Nakayama M, et al. Abnormal DNA methylation, epigenetics, and prostate cancer. *Front Biosci.* 2007;12: 4254–66. PMID: 17485372.
- [35] Hmadcha A, Bedoya FJ, Sobrino F, Pintado E.. Methylation dependent gene silencing induced by interleukin 1 β via nitric oxide production. *J Exp Med.* 1999;190:1595-1604. PMID: 10587350 PMCID: PMC2195731.
- [36] Kwabi-Addo B, Chung W, Shen L, Ittmann M, Wheeler T, Jelinek J, et al. Age-related DNA methylation changes in normal human prostate tissues. *Clin Cancer Res.* 2007;13:3796–3802. PMID: 17606710.
- [37] Bastian PJ, Yegnasubramanian S, Palapattu GS, Rogers CG, Lin X, De Marzo AM, et al.. Molecular biomarker in prostate cancer: the role of CpG island hypermethylation. *Eur Urol.* 2004;46:698-708. PMID: 15548435.
- [38] Adkins RM, Krushkal J, Tylavsky FA, Thomas F. Racial differences in gene-specific DNA methylation levels are present at birth. *Birth Defects Res A Clin Mol Teratol.* 2011;91:728-736. doi: 10.1002/bdra.20770. PMID: 21308978 PMCID: PMC3429933.
- [39] Yardy GW, Brewster SF. Wnt signalling and prostate cancer. *Prostate Cancer Prostatic Dis.* 2005;8:119-126. PMID: 15809669.
- [40] Kypka RM, Waxman J.. Wnt/beta-catenin signalling in prostate cancer. *Nat Rev Urol.* 2012;9:418-428. doi: 10.1038/nrurol.2012.116. PMID: 22710668.

- [41] Sandsmark E, Hansen AF, Selnæs KM, Bertilsson H, Bofin AM, Wright AJ, et al. A novel non canonical Wnt signature for prostate aggressiveness. *Oncotarget*. 2017; 8:9572–9586. . doi: 10.18632/oncotarget.14161. PMID: 28030815
PMCID: PMC5354754.
- [42] Yokoyama NN, Shao S, Hoang BH, Mercola D, Zi X. Wnt signaling in castration-resistant prostate cancer: implications for therapy. *Am J Clin Exp Urol*. 2014;2:27-44. PMID: 25143959 PMCID: PMC4219296.
- [43] Schneider JA, Logan SK. Revisiting the role of Wnt/ β -catenin signaling in prostate cancer. *Mol Cell Endocrinol*. 2017;462(PtA):3-8. doi: 10.1016/j.mce.2017.02.008. PMID: 28189566 PMCID: PMC5550366.
- [44] Miyamoto DT, Zheng Y, Wittner BS, Lee RJ, Zhu H, Broderick KT, et al. RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in anti-androgen resistance. *Science*. 2015;349:1351-1356. doi: 10.1126/science.aab0917. PMID: 26383955 PMCID: PMC4872391.
- [45] Wang L, Dehm SM, Hillman DW, Sicotte H, Tan W, Gormley M, et al. A Prospective Genome-Wide Study of Prostate Cancer Metastases Reveals Association of Wnt Pathway Activation and Increased Cell Cycle Proliferation with Primary Resistance to Abiraterone Acetate-Prednisone. *Ann Oncol*. 2018;29:352-360. doi: 10.1093/annonc/mdx689. PMID: 29069303.
- [46] Cristóbal I, Rojo F, Madoz-Gúrpide J, García-Foncillas J. Cross Talk between Wnt/ β -Catenin and CIP2A/Plk1 Signaling in Prostate Cancer: Promising Therapeutic Implications. *Mol Cell Biol*. 2016;36:1734-1739. doi: 10.1128/MCB.00130-16. PMID: 27090640 PMCID: PMC4907099.

- [47] Katoh M. Canonical and non-canonical WNT signaling in cancer stem cells and their niches: Cellular heterogeneity, omics reprogramming, targeted therapy and tumor plasticity (Review). *Int J Oncol.* 2017;51:1357-1369. doi: 10.3892/ijo.2017.4129. PMID: 29048660 PMCID: PMC5642388.
- [48] Reams RR, Agrawal D, Davis MB, Yoder S, Odedina FT, Kumar N, et al. Microarray comparison of prostate tumor gene expression in African-American and Caucasian American males: A pilot project study. *Infect Agent Cancer.* 2009;4 Suppl 1:S3. doi: 10.1186/1750-9378-4-S1-S3. PMID: 19208208 PMCID: PMC2638462.
- [49] Wang BD, Ceniccola K, Yang Q, Andrawis R, Patel V, Ji Y, et al. Identification and Functional Validation of Reciprocal microRNA-mRNA Pairings in AA Prostate Cancer Disparities. *Clin Cancer Res.* 2015;21:4970-4984. doi: 10.1158/1078-0432.CCR-14-1566. PMID: 26089375 PMCID: PMC4631799.
- [50] Horning AM, Wang Y, Lin CK, Louie AD, Jadhav RR, Hung CN, et al.. Single-cell RNA-seq reveals a subpopulation of prostate cancer cells with enhanced cell cycle-related transcription and attenuated androgen response. *Cancer Res.* 2018;78:853-864. doi: 10.1158/0008-5472.CAN-17-1924. PMID: 29233929.
- [51] Putluri N, Shojaie A, Vasu VT, Nalluri S, Vareed SK, Putluri V, et al. Metabolomic profiling reveals a role for androgen in activating amino acid metabolism and methylation in prostate cancer cells. *PLoS One.* 2011;6:e21417. doi: 10.1371/journal.pone.0021417. PMID: 21789170 PMCID: PMC3138744.
- [52] Pressinotti NC, Klocker H, Schäfer G, Luu VD, Ruschhaupt M, Kuner R, et al.. Differential expression of apoptotic genes PDIA3 and MAP3K5 distinguishes between

low- and high-risk prostate cancer. *Mol Cancer*. 2009;8:130. doi: 10.1186/1476-4598-8-130. PMID: 20035634 PMCID: PMC2807430.

[53] Rigau M, Oliván M, García M, Sequeiros T, Montes M, Colás E, et al. The present and future of prostate cancer urine biomarkers. *Int J Mol Sci*. 2013;14:12620-12649. doi: 10.3390/ijms140612620. PMID: 23774836 PMCID: PMC3709804.

[54] Al-Khallaf H. Isocitrate dehydrogenases in physiology and cancer: biochemical and molecular insight. *Cell Biosci*. 2017;7:37. doi: 10.1186/s13578-017-0165-3. eCollection 2017. PMID: 28785398 PMCID: PMC5543436.

[55] Azmi AS, Bao B, Sarkar FH. Exosomes in cancer development, metastasis, and drug resistance: a comprehensive review. *Cancer Metastasis Rev*. 2013;32:623-642. doi: 10.1007/s10555-013-9441-9 PMID: 23709120 PMCID: PMC3843988.

Figure legends

Figure 1. A. Pictorial representation for the distribution of fusion and mutation associated tumors in different methylation clusters. B. The distribution of African-American and Caucasian patients in different class of methylation clusters. African-American patients are more likely to belong to MC1 whereas Caucasian preferentially belongs to MC3.

Figure 2. A. Pictorial representation of differential regulation of canonical (Wnt/ β -catenin) and non-canonical (Wnt/ Ca^{+2}) Wnt signaling pathways among methylation cluster 1 and 3. B. Wnt (Wnt/ Ca^{+2}) signaling genes were significantly overexpressed in CA cohort as compared to AA cohort from GRID data from GenomeDx.

Figure 3. Epigenetic changes in Caucasian and African-American Prostate cancer patients. A. Heat map representing the methylation variation in CA and AA. B. Heat map represents the methylation variation in benign vs tumor tissues in CA. C. Heat map represents the methylation variation in benign vs tumor tissues in AA. Yellow color correspond to hypermethylated genes in both the groups whereas blue color represents the hypomethylated genes.

Figure 4. Comparison of Kaplan--Meier curve in Caucasian and African American Prostate cancer patients using three prostate cancer databases using survexpress software. A. PRAD-TCGA prostate adenocarcinoma database shows survival days. B. Sboner prostate database survival months. C. Taylor prostate database shows recurrence months.

Horizontal axis represents time to event (days in A, months in B and C). Red and Green curves denote High- and Low-risk groups respectively. The red and green numbers below horizontal axis represent the number of individuals not presenting the event of the corresponding risk group along time. The number of individuals, the number of censored, and the CI of each risk group are shown in the top-right insets. CI represent Concordance index, P represent P value.

Table 1. Clinical profile of Prostate cancer patients.

Patient ID	Race	Gleason Score	PSA	Stage
CA1	Caucasian	4 + 5	15.5	pT3b
CA2	Caucasian	4 + 3	6	pT2c
CA3	Caucasian	5 + 4	1.5	pT3a
AA1	African American	3 + 4	20.7	pT3b
AA2	African American	3 + 4	4.8	pT2c
AA3	African American	3 + 4	10	pT3b

Table 2. Reduced representation bisulfite sequencing data information for prostate tissues (benign and tumor) of African American men and Caucasian men.

Race/ Group	Type of tissue	Raw reads	Unique mapped reads	Mapping percentage
CA1	Benign	12282536	5416647	44.1%
	Tumor	12732930	5803964	45.6%
CA2	Benign	10462111	4675127	44.7%
	Tumor	10607889	4840047	45.6%
CA3	Benign	7160756	3072980	42.9%
	Tumor	11087471	5098151	46.0%
AA1	Benign	12788974	5762829	45.1%
	Tumor	13354714	5831607	43.7%
AA2	Benign	7520582	3363084	44.7%
	Tumor	12431991	5216656	42.0%
AA3	Benign	10402284	4513597	43.4%
	Tumor	9577045	3958823	41.3%

Table 3. Pathways associated with differentially expressed genes in methylation cluster 1 and 3

	Methylation cluster 1 (African-American)		Methylation cluster 3 (Caucasian)	
	Pathways	Genes	Pathways	Genes
Upregulated	Chemokine signaling and Immune cell or inflammation related signaling	MCP-3, Eotaxin, PKC, MLCP, CCR3, PI3K, cPLA2, Gβ, AIM2, CASP1, CASP5, VAV2, PXN, VCAM1, PRKCQ, NOX3, MMP16, CLDN18, PRKCZ, CTNNA2, IRS2, DLC1, MMP17, ITK, PRKCB	Amino Acid Metabolism	BCAT1, BCAT2, HMGCLL1, EHHADH, ALDH2, ALDH1A2, ALDH1L1, ALDH3A1
	PI3K Signaling	VAV2, PLCZ1, PTPRC, IRS2, PLCL2, PRKCZ, NFATC1, PPP3CA, PRKCB		
Downregulated	Cell Cycle related pathways	CDKN2A, HDAC9, CCND2, CCNA1, CCND3, PPP2R2A, PPP2R3A, HDAC7, CUL1, CDK6, HDAC10, PPP2R5C, CDKN1B, NRG1, SMAD3	Cytokine and Chemokine signaling	VCAM1, IL-6, AKT, PI3K, JAK1, JAK3, PLCγ1, mTOR, NFAT, MHC-II, MCP-3, SDF-1, CAMK, PI3Kγ, MLCP, CCR3,
			PI3K Signaling	NFAT, TAPP, PLCγ1, CAMK, IP3R, Calcineurin, TLR4, PI3K, AKT

Table 4. Comparison of hypomethylated genes in African American (hypermethylated in Caucasian) in RRBS analysis and hypomethylated in TCGA dataset (Methylation mean cut off < 0.2).

Gene	RRBS		TCGA data	
	P value	Methylation difference	P value	Methylation Mean
RPS29	0.002	-0.23	1.07E-09	0.04
ZNF180	0.03	-0.15	1.04E-02	0.03
POLD4	0.04	-0.14	3.58E-06	0.14
AKIRIN2	0.01	-0.13	4.67E-15	0.03
SLC41A3	0.02	-0.13	4.69E-03	0.06
PDIA3	0.01	-0.12	7.33E-02	0.12
CPSF2	0.04	-0.11	1.62E-07	0.04
SLC35C2	0.01	-0.10	5.60E-08	0.05

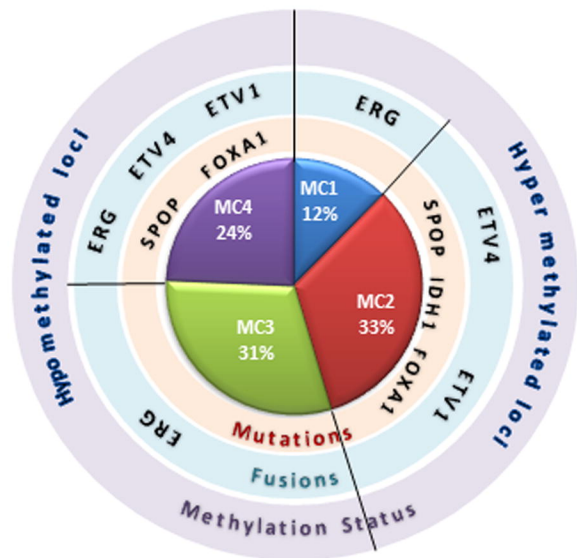
Table 5. Comparison of hypermethylated genes in African American (hypomethylated in Caucasian) in RRBS analysis and hypermethylated in TCGA dataset (Methylation mean cut-off > 0.3).

Gene	RRBS		TCGA data	
	P value	Methylation difference	P value	Methylation Mean
IDH3B	0.035	0.262	9.95E-09	10.38
LRP10	0.017	0.105	1.66E-04	12.16
BLVRB	0.039	0.166	7.89E-03	10.45
ZNF256	0.021	0.135	1.02E-05	6.71
LIME1	0.047	0.205	3.03E-07	6.34
CD72	0.038	0.106	4.15E-15	4.47
URM1	0.017	0.266	1.49E-03	10.25
PSAP	0.008	0.467	5.37E-09	14.88
PIKFYVE	0.046	0.106	1.12E-06	9.96
ARL6IP1	0.003	0.202	3.14E-15	12.12
BTBD2	0.032	0.130	6.64E-07	11.49

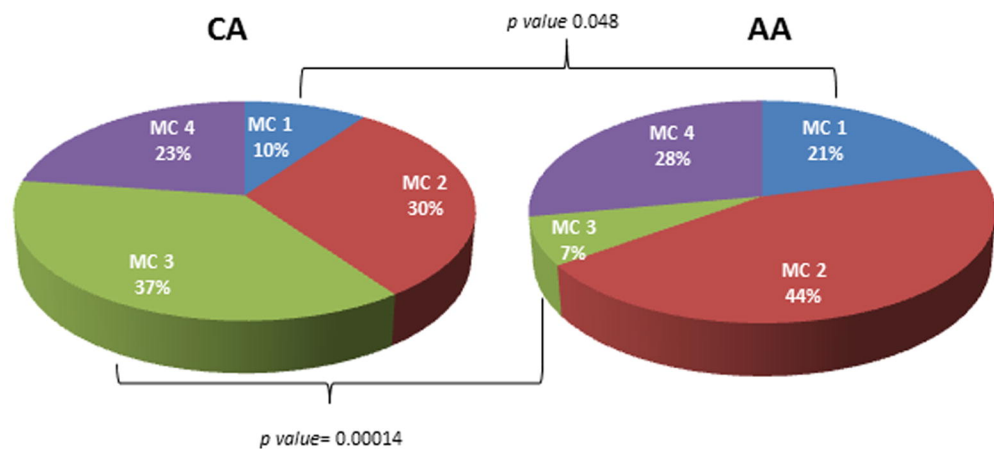
Table 6. Evaluation of survival analysis and prognostic markers in three different prostate cancer databases

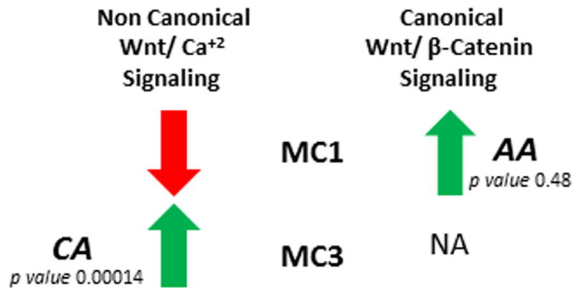
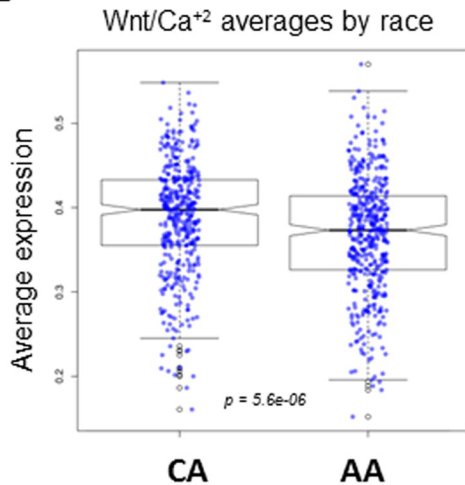
Datasets	Caucasian			African American		
	Risk group p Value	DEG between risk groups	Genes	Risk group p Value	DEG between risk groups	Genes
PRAD-TCGA database [26]	0.998	14	HOXD12, C6orf108, C17orf46, SIVA1, OLFM1, TMEM14A, C2orf56, BLVRB, CA9, ZNF566, ARL6IP1, CCDC104, MBD1, PSAP	0.0474	12	KLHL12, EZH1, RPS29, ALDH3B1, LCN1, MTHFD1, GALNT6, AKIRIN2, ODF3L1, NDUFB1, SLC35C2, GPR37
Sboner database [27]	0.00013	10	ATP5G2, C6orf108, CD72, SIVA1, OLFM1, CEBPZ, BLVRB, CA9, IDH3B, PSAP	P = 0.02	7	EZH1, ALDH3B1, LCN1, MTHFD1, POLD4, PDIA3, GPR37
Taylor database [28]	6.01e-06	10	ATP5G2, SLC34A3, PIKFYVE, ATXN1L, DYNC1I2, CEBPZ, CA9, ARL6IP1, URM1, MBD1	P = 0.0077	8	EZH1, LCN1, GALNT6, POLD4, C4orf39, SLC41A3, ODF3L1, CPSF2

A

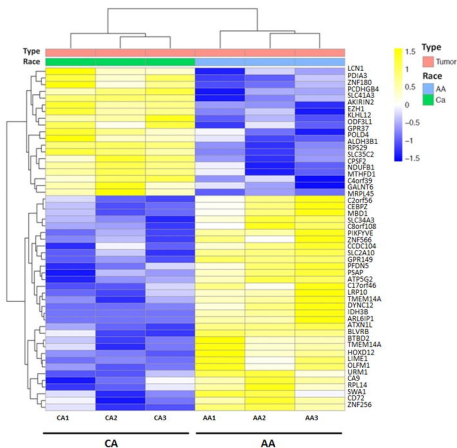


B

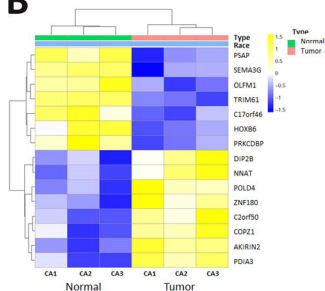


A**B**

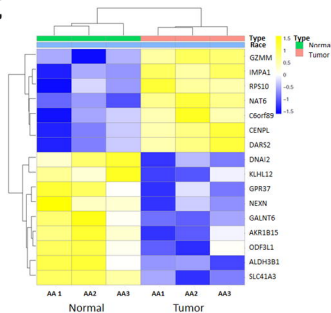
A



B



C



Caucasian
(Tumor vs Benign)

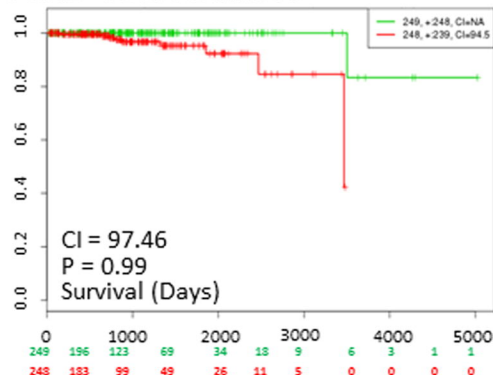
- Lipid Antigen presentation by CD1
- Glutathione Redox Reaction
- Lactose degradation
- EIF4 signaling

African American
(Tumor vs Benign)

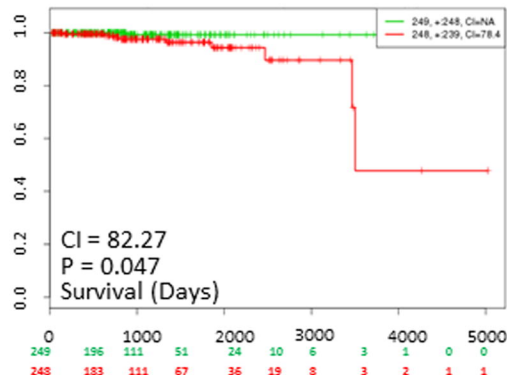
- Histamine degradation
- Fatty acid oxidation
- Tryptophan degradation
- Putrescine Degradation III

CA

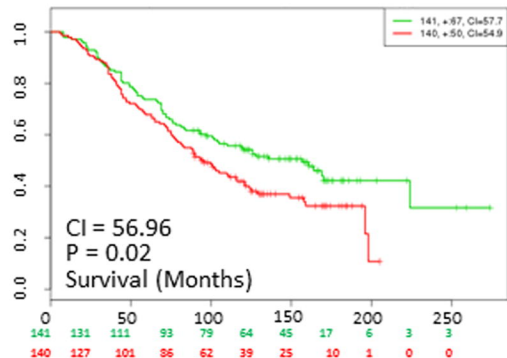
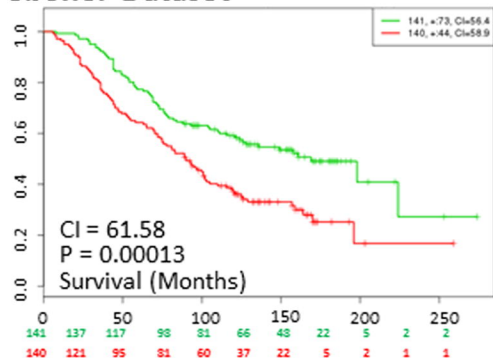
A PRAD-TCGA Dataset



AA



B Sboner Dataset



C Taylor Dataset

