Published in final edited form as:

Surgery. 2021 October ; 170(4): 1160–1167. doi:10.1016/j.surg.2021.03.031.

Genomics of Black American Colon Cancer Disparities: An RNA-Seq Study from an Academic, Tertiary Referral Center

Ivy N. Haskins, MD^{1,2}, Bi-Dar Wang, PhD³, James P. Bernot, MS³, Edmund Cauley, MPhil³, Anelia Horvath, PhD³, John H Marks, MD⁵, Norman H. Lee, PhD^{3,*}, Samir Agarwal, MD FACS FASCRS^{1,4,*}

¹Department of Surgery, George Washington University, Washington, DC

²Department of Surgery, University of Nebraska Medical Center, Omaha, NE

³Department of Pharmacology & Physiology, School of Medicine and Health Sciences, GW Cancer Center, George Washington University, Washington, DC

⁴Department of Colorectal Surgery, Cleveland Clinic Florida, Weston, FL

⁵Division of Colorectal Surgery, Lankenau Medical Center, Wynnewood PA

Abstract

Introduction: Black Americans (BA) have a higher incidence and mortality rate from colorectal cancer compared to their Non-Hispanic White American (NHW) counterparts. Even when controlling for sociodemographic differences between these two populations, BA remain disproportionately affected by colorectal cancer. The purpose of our study was to determine if differences in gene expression between BA and NHW colon cancer specimens could help explain differences in the incidence and mortality rate between these two populations.

Methods: BA and NHW undergoing colon resection for stage I, II, or III colon cancer at a single institution were identified. BA and NHW patients were matched for age, gender, and colon cancer stage to minimize the risk of confounding variables. Tissue samples were obtained at the time of colon resection and were analyzed using RNA sequencing (RNA-Seq) to determine if there were differences in the expression of genes and biologic processes between the two groups.

Results: A total of 17 colon cancer specimens were analyzed; 8 (47.1%) patients were BA. A total of 456 genes were identified as being expressed differently (i.e. up- or down-regulated) in BA compared to NHW colon cancer specimens. Moreover, 500 different genetic pathways were noted to be significantly over-represented with differentially expressed genes in our comparison of BA and NHW colon cancer specimens, the majority of which plays a role in inflammation and immune cell function.

Conclusions: Significant differences in gene expression and genetic pathways exist between BA and NHW. Additional and multi-institutional and registry-based studies are needed to validate

* **Corresponding Authors:** Samir Agarwal, MD FACS FASCRS, Department of Colorectal Surgery, Cleveland Clinic Florida, Mail Code WSTN, 2950 Cleveland Clinic Boulevard, Weston, FL 33331, Phone 954.659.5278, Fax 954.659.5252, agarwas2@ccf.org, Norman H. Lee, PhD, Department of Pharmacology & Physiology, The George Washington University, 2300 I Street N.W., Washington, DC 20037, Phone: (202) 994-8855, Fax: (202) 994-2870, nhlee@gwu.edu.

our findings and to further elucidate the contribution that these differences have to the overall incidence and mortality rate from colon cancer in these two patient populations.

ARTICLE SUMMARY

The incidence and mortality rate from colon cancer varies significantly between Black Americans (BA) and Non-Hispanic White Americans (NHW). This study used RNA-Seq to identify difference in expression of 456 genes and 500 genetic pathways between BA and NHW colon cancer specimens, which may help to explain differences in the progression of colon cancer between these two populations.

INTRODUCTION

Colorectal cancer is the third leading cause of cancer-related death in the United States and worldwide.¹⁻³ According to the American Cancer Society, it is estimated that 104,610 Americans will be diagnosed with colon cancer and that 53,200 Americans will die from colon cancer this year.³ Unfortunately, the incidence and mortality rate associated with colon cancer is not evenly distributed across racial and ethnic groups in the United States.³⁻⁴ In fact, the incidence of colon cancer in black Americans (BA) is 20% higher than that of non-Hispanic white Americans (NHW).⁴ Furthermore, the mortality rate from colon cancer in BA is 40% higher than that of NHW.⁴

Currently, the preponderance of medical literature addressing differences in cancer outcomes between BA and NHW cite sociodemographic features, including lifestyle and healthcare differences between BA and NHW.⁵⁻⁶ When these lifestyle and healthcare differences between BA and NHW are used to explain differences in cancer outcomes, it is often concluded that these differences result in later presentation, lack of follow-up, and more advanced-stage cancer at the time of diagnosis in the BA population.⁵⁻⁶ In order to circumvent potential sociodemographic differences in the BA and NHW, Lai et al performed a matched analysis of BA and NHW patients diagnosed with colon cancer based on demographic data and colon cancer stage and still found differences in overall two-, three- and five-year survival.⁶ These findings suggest that intrinsic biological and genetic factors in colon cancer tumors may account for some of the disparate cancer outcomes and survival between BA and NHW.

The search for genomic and genetic factors contributing to colon cancer disparities has led to the identification of differential mutations, chromosomal aberrations, and DNA methylation between BA and NHW populations in genes responsible for both the development and progression of cancer.⁷⁻¹¹ Nevertheless, in the three most commonly cited articles addressing colorectal cancer gene sequencing, BA cases are under-represented.⁹ Furthermore, RNA sequencing (RNA-Seq) to identify differences in genomic and genetic factors between BA and NHW has been under-utilized. For example, the most cited genomic study in colorectal cancer used Affymetrix arrays, which is a hybridization-based methodology.¹² The advantages of the RNA-Seq approach over hybridization-based strategies include greater dynamic range of measured gene expression levels as well as the ability to categorize more differentially expressed genes.¹³⁻¹⁵ Therefore, the purpose of our study was to better elucidate the differences in gene expression between BA and NHW

using RNA-Seq. As a pilot study, we hypothesize that the identification of race-specific differentially expressed genes can be leveraged as potential biomarkers for prognostication or candidate genes to understand the molecular basis of racial cancer disparities. Moreover, race-specific gene expression differences in colon cancer are hypothesized to be applicable in other cancers, which can be further explored in larger multi-institutional or registry-based studies.

METHODS

Inclusion and Exclusion Criteria

Adult BA and NHPA patients aged 18 years or older who were diagnosed with colon cancer by screening colonoscopy at an academic, tertiary referral hospital from July 2013 through April 2016 were prospectively identified. A diagnosis of colon cancer by screening colonoscopy is confirmed based on pathology from the procedure. Patients with a confirmed diagnosis of colon cancer are then referred to one of two colorectal surgeons. As part of their evaluation, patients undergo computed tomography (CT) scans of the chest, abdomen, and pelvis to rule out metastatic disease and for evaluation of lymphadenopathy. Based on the results of these tests, patients are given a preoperative, clinical colon cancer stage. For the purposes of our study, all adult (> 18 years of age) patients undergoing elective colon resection for clinical stages I, II, or III colon cancer without a previous diagnosis of cancer and who had not undergone neoadjuvant therapy for their colon cancer were eligible for study inclusion. Pediatric patients (< 18 years of age), patients undergoing emergency colon resection for their colon cancer (i.e. large bowel obstruction or colonic perforation), patients with clinical stage IV colon cancer, patients who did not undergo colon resection, patients with a final pathologic stage other than stages I, II, or III, patients with a history of any previous cancer, patients who underwent neoadjuvant therapy for their colon cancer, and those patients who were unable to make their own medical decisions were excluded from our analysis.

As part of the informed consent process, all patients who were offered study inclusion were informed that their participation in this study would not affect the care that they received for their colon cancer. This study was approved by our Institutional Review Board (IRB # 031341).

Sample Collection and Preservation

Patients enrolled in this study were re-identified the morning of surgery in the preoperative area. A copy of their study consent was included with their surgical consent form. Once the colectomy specimen was removed from the patient, a member of the study team retrieved the study specimen from the operating room. We aimed for a specimen size of at least one cubic centimeter to ensure adequate tissue for analysis. The study specimen was placed in RNeasyTM preservation solution (ThermoFisher Scientific) and immediately stored at -80°C.

Gene Expression Profiling by RNA-Seq

Prior to RNA-Seq of specimens, a pathological staging of I, II, or III was confirmed. The specimens were also matched for age, stage, gender, location of primary tumor, tumor grade, and lymphovascular invasion using Fisher's exact test and a p-value < 0.05 to determine statistical significance to minimize the effect of any confounding variables associated with differences in the genomics of colon cancer between BA and NHPA. RNA extraction, library construction, RNA-Seq, and gene expression analysis were performed as previously described by Maynard et al.¹⁵

Significant differential expression analysis was performed using two strategies in parallel: the CuffDiff utility of Cufflinks that contains normalization modules supporting across-sample comparisons and the outlier-resistant EdgeR-robust package.^{16–18} In addition, ANOVA was applied to define significant differentially regulated transcripts, and p-values were corrected for multiple testing using False Discovery Rate (FDR).¹⁹ Unless otherwise noted, a FDR of 10% (q-value = 0.10) was considered significant for all analyses. Gene expression differences between BA and NHPA were reported as log₂ expression ratio values (BA/NHPA), with a positive log₂ value correlating to higher expression in BA.

Once differences in gene expression between BA and NHPA were identified within our colon cancer specimens, we sought to compare our findings to currently available gene expression differences in other cancers. In order to do this, RNA-Seq data was downloaded from The Cancer Genome Atlas via the Genomic Data Commons portal (<https://gdc.cancer.gov>). Differences in gene expression between BA and NHPA patients was available for comparison for bladder cancer, head and neck squamous cell carcinoma, renal cell carcinoma, lung adenocarcinoma, and uterine corpus endometrial cancer.

Functional and Enrichment Analyses of RNA-Seq

Differences in gene expression between BA and NHPA colon cancer were further assessed for pathway/network enrichment using the Ingenuity Pathway Analysis (IPA) computational platform (Qiagen). This software identifies canonical cell signaling pathways, gene networks, disease categories and biological processes that are significantly over-represented by differentially expressed genes between BA and NHPA. Over-representation is presumed to have biological relevance with identified signaling pathways potentially linked to colon cancer racial disparities. A Fisher's exact test with a p-value < 0.05 was used to define significant over-representation. The magnitude of difference in over-representation of the signaling pathways and biological processes is determined by the $-\log_{10}[\text{p-value}]$, with a threshold of > 1.3 determining statistical significance.²⁰

RESULTS

A total of 25 patients were enrolled in our study; 12 (48%) BA and 13 (52%) NHPA. One patient enrolled in our study was excluded due to a small tumor size (in order to ensure adequate cancer analysis) and one patient was excluded due to final pathology consistent with high-grade, adenomatous polyp. After RNA quality control evaluation, 5 (22%) additional patients and their specimens were excluded, leaving 17 (68%) patients and

their corresponding tumor specimens available for RNA-Seq analysis. Of the 17 patients and their corresponding tumor specimens available for RNA-Seq analysis, 8 (47%) were BA and 9 (53%) were NHWA. As shown in Table 1, these patients were well-matched for gender ($p = 1.00$, two-tailed), age ($p = 1.00$, two-tailed), colon cancer stage ($p = 0.37$, two-tailed), location of primary tumor ($p = 0.33$, two-tailed), tumor grade ($p = 1.00$, two-tailed), and lymphovascular invasion ($p = 0.35$, two-tailed).

RNA-Seq analysis identified 14,388 unique expressed genes with an average FPKM value 1 in either BA or NHWA colon cancer specimens (Supplemental Table 1), indicating that a large fraction of the human genome is expressed in colon cancer. Prior RNA-Seq studies in colon cancer have identified between 7,744 and 10,255 expressed genes.^{20–21} The ability of our study to identify more expressed genes is likely a consequence of the higher sequencing depth (>200 million total reads per sample) compared to the earlier studies that were performed at a depth ranging from ~50 to 90 million reads per sample.^{21–22}

Inspection of our transcriptome data revealed a total of 456 differentially expressed genes at an FDR q -value 0.10 in BA versus NHWA colon cancer (Supplemental Table 2). Of these 456 genes, 173 genes were up-regulated and 283 genes were down-regulated in BA colon cancer specimens relative to the NHWA colon cancer specimens (Supplemental Table 3). The top 20 up-regulated and down-regulated genes based on both the magnitude of differential regulation and q -value are shown in Tables 2 and 3, respectively. The three most differentially expressed genes (up- or down-regulated) between BA and NHWA were SOX30 ($\log_2$ ratio = 13.8, q -value = 0.0006), TRPM3 ($\log_2$ ratio = 10.8, q -value = 0.0002), and RNF38 ($\log_2$ ratio = -11.5, q -value = 0.0002) (Tables 2 and 3). When compared to The Cancer Genome Atlas, SOX30 was also found to be more highly expressed in BA compared to NHWA in renal cell carcinoma and uterine corpus endometrial carcinoma and RNF38 was found to be more highly expressed in NHWA compared to BA in lung squamous cell carcinoma (Table 4). Interestingly, there was no race-specific difference in expression of TRPM3 for any of the five available cancers (Table 4).

The pathways with the most significant difference in expression between BA and NHWA colon cancer are displayed in Figure 1. Differentially expressed genes of known function were significantly enriched in BA in a number of canonical signaling pathways, including agranulocyte adhesion and extravasation ($-\log_{10} p = 8.89$), granulocyte adhesion and extravasation ($-\log_{10} p = 5.47$), hepatic fibrosis ($-\log_{10} p = 3.95$), atherosclerosis signaling ($-\log_{10} p = 3.42$), insulin growth factor (IGF)-1 signaling ($-\log_{10} p = 2.63$), virus entry via endocytic pathways ($-\log_{10} p = 2.56$), tight junction signaling ($-\log_{10} p = 2.54$), and antigen presentation pathway ($-\log_{10} p = 2.33$). On the other hand, the activity of several pathways in immune cell function, including THOP1 signaling ($-\log_{10} p = 5.21$), complement system ($-\log_{10} p = 4.40$), acute phase response signaling ($-\log_{10} p = 3.49$), and leukocyte extravasation signaling ($-\log_{10} p = 2.60$) were predicted by IPA to be suppressed in BA colon cancer relative to their NHWA counterparts (Figure 1). A complete listing of pathways differentially expressed in BA versus NHWA is available in Supplementary Table 3.

A total of 500 different disease and biological processes, which included 456 genes, were identified as differentially expressed in BA versus NHWA. A full listing of these differences is available in Supplementary Table 4. Approximately one-third of the 456 differentially expressed genes could be assigned to categories associated with inflammation and immune cell function (Figure 2). Particularly noteworthy, IPA analysis predicted suppression of numerous immune cell processes, including cell movement and migration ($-\log_{10} p = 16.83$), development of vasculature ($-\log_{10} p = 11.48$), angiogenesis ($-\log_{10} p = 8.97$), recruitment of leukocytes ($-\log_{10} p = 8.77$), cell survival ($-\log_{10} p = 8.89$), attraction of immune cells ($-\log_{10} p = 5.04$), chemotaxis of immune cells ($-\log_{10} p = 4.79$), migration of phagocytes ($-\log_{10} p = 4.72$), homing of leukocytes ($-\log_{10} p = 4.58$), and recruitment of lymphocytes ($-\log_{10} p = 4.24$) in BA colon cancer relative to NHWA colon cancer (Figure 2).

DISCUSSION

Explanations for racial and ethnic differences in the incidence of, and mortality from, colon cancer are complex. According to the Surveillance, Epidemiology, and End Results (SEER) Program, the death-to-diagnosis ratio for colon cancer for NHWA is 36/100 persons and a death-to-diagnosis ratio for colon cancer for BA is 42/100 persons.²³ While the overall incidence of colon cancer continues to decline and the overall survival from colon cancer continues to improve, differences in the proportion of NHWA and BA affected by colon cancer continues to persist.²⁴ Using a matched analysis to control for potential sociodemographic and cancer pathology contributors to these disparities, we found significant differences in gene expression patterns as well as differences in predicted activities of signaling pathways and biological processes between BA and NHWA colon cancer specimens. The findings from our study are important because the three genes that we identified as having the greatest differences in expression between BA and NHWA have not previously been discussed in colon cancer. These genes and their contribution to the differences in colon cancer incidence and progression between BA and NHWA warrants further investigation.

Of the 456 differentially expressed genes identified in our RNA-Seq-based study, SOX30 exhibited the greatest over-expression in BA compared to NHWA colon cancer (Table 2). SOX refers to a group of genes that contain the evolutionarily conserved high-mobility group (HMG) box from SRY, a gene involved in sex determination.²³ The SOX genes encode transcription factors that have been associated with tumorigenesis, cancer progression, and cancer metastasis.²³ The family of SOX genes includes more than 20 members, which are divided into eight groups: SoxA through SoxH.²³ The HMG portion of each SOX gene is responsible for DNA binding involved in cell development, which can be either up-regulated or down-regulated in cancer cells.²³ Interestingly, up-regulation of a particular SOX gene may lead to cancer in one organ while down-regulation of the same SOX gene may lead to cancer in a different organ, thereby exhibiting cell context-specific effects.²³ SOX30 is the only gene in the SoxH group and it is associated with tumor suppression and apoptosis through the transcriptional activation of the tumor suppressor gene *p53*.²³ Interestingly, increased expression of SOX30 is associated with a better prognosis in patients with lung adenocarcinoma and prostate cancer.^{23,25} In terms

of BA vs NHPA, increased levels of the SOX30 protein have been found in BA with papillary renal cell carcinoma.²⁶ Further studies are needed to determine if the up-regulation of SOX30 in BA colon cancer patients offers a survival benefit or if it is associated with increased tumor malignancy.

TRPM3 was the second most up-regulated gene in BA colon cancer relative to NHPA colon cancer (Table 2). This gene belongs to the melastin transient receptor potential (TRPM) family. The TRPM family, which is comprised of eight different cation channels, M1-M8, has been shown to mobilize Ca²⁺ via store-operated Ca²⁺ entry (SOCE).²³ Calcium signaling plays an important role in cell homeostasis.²³ Following disruption of homeostatic intracellular Ca²⁺ levels, cell proliferation increases and resistance to apoptosis ensues.^{28–29} Up- and down-regulation of the TRPM family has been associated with the prognosis for several cancers, including breast, colorectal, gastric, liver, melanoma, and prostate.³⁰ Interestingly, TRPM1, TRPM2, TRPM4, TRPM5, and TRPM6 have been shown to be differentially expressed in colon carcinoma compared to control tissues.^{27,30} To our knowledge, however, no study has identified differences in TRPM3 expression as impacting the clinical course of colon cancer. Furthermore, no study has investigated differences in expression of the TRPM family in BA versus NHPA. Based on our results, it is possible that the up-regulation of TRPM3 does have clinical implications in colon cancer. Therefore, additional studies are needed to determine if differential expression of TRPM3 impacts the overall survival of BA and NHPA patients with colon cancer.

The most down-regulated gene in BA colon cancer relative to NHPA colon cancer was ring finger protein 38 (RNF38). RNF38 is found on the short arm of chromosome 9 in an area that is associated with many cancers.³¹ RNF38 participates in the ubiquitination of p53.³¹ Ubiquitination plays an important role in many normal cell processes as well as in the development and progression of cancer.^{31–33} In a recent study by Peng et al., overexpression of RNF38 mRNA and protein levels was associated with a worse prognosis of patients with hepatocellular carcinoma.³³ Nevertheless, there are currently no studies investigating differences in BA versus NHPA or that identify a role for RNF38 in the development and progression of colon cancer. The role of RNF38 in colon cancer and the impact of its down-regulation warrants further investigation.

Perhaps the most intriguing aspect of our study was the cumulative gene expression differences between BA and NHPA colon cancer that are computationally predicted to lead to differential activity status of canonical signaling pathways and biological processes (Figures 1 and 2). Of particular interest is the over-expression of agranulocyte and granulocyte adhesion and diapedesis cytokine genes (e.g. CCL2, CCL19, CCL21, CCL25, CCL26 CXCL1, CXCL2, CXCL3, CXCL12, CXCL13), which portend suppression of immune cell function and recruitment to the tumor microenvironment of BAs compared to their NHPA counterparts.³⁴ Recruitment of leukocytes, attraction of immune cells, chemotaxis of immune cells, migration of phagocytes, homing of leukocytes, migration of myeloid cells, and recruitment of lymphocytes were all predicted to be suppressed in BA colon cancer specimens. Two previous studies have shown differences in colorectal cancer tumor microenvironments between BA and NHPA, with BA tumor microenvironments being more pro-inflammatory with a loss of immune response and cell repair.^{4,35} Our RNA-

Seq data and pathway analyses, which predicted suppressed recruitment of lymphocytes in BA specimens, are in agreement with the findings from previous studies. It will be of interest in future studies to determine the profile of other intratumoral-infiltrating immune cells in BA versus NHPA colon cancer given the role of immunotherapy for metastatic colon cancer.^{36–37}

To our knowledge, our study introduces the greatest number of gene expression differences in BA and NHPA colon cancer, to date. Furthermore, it highlights that even though differences in expression of SOX30, TRPM3, and RNF38 have been found in other cancers, there is currently a paucity in all of the cancer literature that addresses differences in expression between BA and NHPA. Nevertheless, despite our findings, our study does have some limitations, which are worth discussing. First, while we identify a large number of differences in gene expression, the clinical implication that these differences have on the prognosis of colon cancer remains to be determined. Currently, there is a paucity of publicly available BA genomic and clinical data in the form of exome sequencing, RNA-Seq, and survival information.³⁸ Moreover, mRNA and protein level correlations are poor and vary across cancers.³⁹ Consequently, candidate genes identified in this study will require future functional validation. Second, this is a single-institution study with a small sample size. While this study was performed at a tertiary referral center, there are several competing hospital systems in close geographic proximity to our institution. Additionally, we were unable to actively promote or advertise this study due to limited funding. Furthermore, there was an initial learning curve associated with preservation and amplification of our specimens, which further contributed to the small sample size. Next, the migration patterns of BA create the potential for sub-group populations with different gene expressions.⁹ Therefore, our findings may only be representative of BA gene expression in the northeast United States. Finally, our study is based on racial/ethnicity self-identification as opposed to ancestry analysis predicated on ancestry informative markers (AIMs), where race/ethnicity information and genetic ancestry computed using AIMs may not always agree.⁴⁰ Therefore, future RNA-Seq studies involving disparities research should employ both self-identification and AIMs information. Despite these limitations, however, this study was intended to be a pilot study that would identify differences in expression of genes and genetic pathways between BA and NHPA that could be further investigated in multi-institutional or registry-based trials and we believe that we have achieved this goal.

CONCLUSION

Racial disparities in colon cancer incidence and survival are likely due to an interplay of sociodemographic factors as well as differences in tumor biology.⁴ Using a matched analysis and RNA-Seq, we were able to identify significant differences in the expression of 456 genes and significant over-representation of differentially expressed genes in 500 genetic pathways between BA and NHPA colon cancer specimens. Further studies are needed to validate our findings and to elucidate the role that these differences in gene and pathway expression have in the development, progression, and disease-free and overall survival of colon cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Claire Fraser and the University of Maryland Institute for Genome Sciences for performing the RNA-Seq.

COI/DISCLOSURE AND FUNDING/SUPPORT

This research was supported by a Katzen Cancer Research Center Award Grant from the GW Cancer Center (S.A. and N.H.L.) and NCI grant R01-CA204806 (N.H.L.).

REFERENCES

1. Fan J, Li J, Guo S, Tao C, et al. Genome-Wide DNA Methylation Profiles of Low- and High-Grade Adenoma Reveals Potential Biomarkers for Early Detection of Colorectal Carcinoma. *Clin Epigenetics*. 2020; 21.1: 56.
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2019. *CA Cancer J Clin*. 2019; 69.1: 7–34. [PubMed: 30620402]
3. Cancer Facts & Figures 2020. <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2020.html>. Accessed May 9, 2020.
4. Colorectal Cancer Facts & Figures 2020–2022. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/colorectal-cancer-facts-and-figures/colorectal-cancer-facts-and-figures-2020-2022.pdf>. Accessed May 9, 2020.
5. Esnaola NF, Ford ME. Racial Differences and Disparities in Cancer Care and Outcomes: Where's the Rub? *Surg Oncol Clin N Am*. 2012; 21.3: 417–437. [PubMed: 22583991]
6. Lai Y, Wang C, Civan JM, Palazzo JP, et al. Effects of Cancer Stage and Treatment Differences on Racial Disparities in Survival from Colon Cancer: A United States Population-Based Study. *Gastroenterology*. 2016; 150: 1135–1146. [PubMed: 26836586]
7. Ashktorab H, Daremipouran M, Devaney J, Varma S, et al. Identification of Novel Mutations by Exome Sequencing in African American Colorectal Cancer Patients. *Cancer*. 2015; 121.1: 34–42. [PubMed: 25250560]
8. Brim H and Ashktorab H. Genomics of Colorectal Cancer in African Americans. *Next Gener Seq Appl*. 2016; 3.2: 133, DOI: 10.4172/2469-9853.1000133. [PubMed: 27917406]
9. Guda K, Veigl M, Varadan V, Nosrati A, et al. Novel Recurrently Mutated Genes in African American Colon Cancers. *Proc Natl Acad Sci U S A*. 2015; 112.4: 1149–1154. [PubMed: 25583493]
10. Wang Z, Li L, Guda K, Chen Z, et al. Adverse Clinical Outcomes Associated with Mutations that Typify African American Colorectal Cancers. *J Natl Cancer Inst*. 2016; 108.12: djw164. [PubMed: 27582379]
11. Augustus GJ and Ellis NA. Colorectal Cancer Disparity in African Americans: Risk Factors and Carcinogenic Mechanisms. *Am J Pathol*. 2018; 118.2: 291–303.
12. The Cancer Genome Atlas Network. Comprehensive Molecular Characterization of Human Colon and Rectal Cancer. *Nature*. 2012; 487: 330–337. [PubMed: 22810696]
13. Wang C, Gong B, Bushel PR, Thierry-Mieg J, et al. The Concordance between RNA-Seq and Microarray Data Depends on Chemical Treatment and Transcript Abundance. *Nat Biotechnol*. 2014; 32.9: 926–932. [PubMed: 25150839]
14. Wang Z, Gerstein M, Snyder M. RNA-Seq: A Revolutionary Tool for Transcriptomics. *Nat Rev Genet*. 2009; 10.1: 57–63. [PubMed: 19015660]
15. Maynard TM, Horvath A, Bernot JP, Karpinski BA, et al. Transcriptional Dysregulation in Developing Trigeminal Sensory Neurons in the LgDel Mouse Model of DiGeorge 22q11.2 Deletion Syndrome. *Human Molecular Genetics*. 2020; 29: 9.1: 1580.

16. Kim D, Langmead B, Salzberg SL. HISAT: A Fast Spliced Aligner with Low Memory Requirements. *Nat Methods*. 2015; 1.4: 357–360.
17. McCarthy DJ, Chen Y, Smyth GK. Differential Expression Analysis of Multifactor RNA-Seq Experiments with Respect to Biological Variation. *Nucleic Acids Res*. 2012; 40.10: 4288–4297. [PubMed: 22287627]
18. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics*. 2010; 26.1: 139–140. [PubMed: 19910308]
19. Benjamini Y and Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Statist Soc B*. 1995; 57.1: 289–300.
20. Lu TP, Tsai MH, Lee JM, Hsu CP, et al. Identification of a Novel Biomarker, SEMA5A, for Non-Small cell Lung Carcinoma in Nonsmoking Women. *Cancer Epidemiol Biomarkers Prev*. 2010; 19.10: 2590–2597. [PubMed: 20802022]
21. Lee JR, Kwon CH, Cho Y, Park HJ, et al. Transcriptome Analysis of Paired Primary Colorectal Carcinoma and Liver Metastases Reveals Fusion Transcripts and Similar Gene Expression Profiles in Primary Carcinoma and Liver Metastases. *BMC Cancer*. 2016; 16: 539. DOI: 10.1186/s12885-016-2596-3. [PubMed: 27461012]
22. Xu L, Wang R, Ziegelbauer J, Wu WW, et al. Transcriptome Analysis of Human Colorectal Cancer Biopsies Reveals Extensive Expression Correlations among Genes Related to Cell Proliferation, Lipid Metabolism, Immune Response, and Collagen Catabolism. *Oncotarget*. 2017; 43: 74703–74719.
23. Grimm D, Bauer J, Wise P, Kruger M, et al. The Role of SOX Family Members in Solid Tumours and Metastasis. *Semin Cancer Biol*. 2019. DOI: 10.1016/j.semcancer.2019.03.004.
24. Yothers G, Sargent DJ, Wolmark N, Goldberg RM, et al. Outcomes Among Black Patients with Stage II and III Colon Cancer Receiving Chemotherapy: An Analysis of ACCENT Adjuvant Trials. *J Natl Cancer Inst*. 2011; 103.20: 1498–1506. [PubMed: 21997132]
25. Fu Q, Sun S, Yang F, Mao T, et al. SOX30, a Target Gene of miR-653–5p, Represses the Proliferation and Invasion of Prostate Cancer Cells through Inhibition of Wnt/ β -Catenin Signaling. *Cell Mol Biol Lett*. 2019; 24: 71. DOI: 10.1186/s11658-019-0195-4. [PubMed: 31889959]
26. Paulucci DJ, Sfakianos JP, Skanderup AJ, Khan K, et al. Genomic Differences between Black and white Patients Implicate a Distinct Immune Response to Papillary Renal Cell Carcinoma. *Oncotarget*. 2017; 8.3: 5196–5205. [PubMed: 28029648]
27. Perez-Riesgo E, Gutierrez LG, Ubierna D, Acedo A, et al. Transcriptomic Analysis of Calcium Remodeling in Colorectal Cancer. *Int J Mol Sci*. 2017; 18.5: E922. [PubMed: 28448473]
28. Villalobos C, Sobradillo D, Hernandez-Morales M, Nunez L. Calcium Remodeling in Colorectal Cancer. *Biochim Biophys Acta Mol Cell Res*. 2017; 1864.6: 843–849. [PubMed: 28087343]
29. Prevarskaya N, Ouadid-Ahidouch H, Skryma R, Shuba Y. Remodelling of Ca²⁺ Transport in Cancer: How it Contributes to Cancer Hallmarks? *Philos Trans R Soc Lond B Biol Sci*. 2014; 369.1638: 20130097. [PubMed: 24493745]
30. Qin F, Loa L, Huang M, Tan H, et al. Evaluation of the TRPM Protein Family as Potential Biomarkers for Various Types of Human Cancer using Public Database Analyses. *Exp Ther Med*. 2020; 20.2: 770–785. [PubMed: 32742323]
31. Sheren JE, Kassenbrock CK. RNF38 Encodes a Nuclear Ubiquitin Protein Ligase that Modifies p53. *Biochem Biophys Res Commun*. 2013; 440.4: 473–478. [PubMed: 23973461]
32. Popovic D, Vucic D, Dikic I. Ubiquitination in Disease Pathogenesis and Treatment. *Nat Med*. 2014; 20.11: 1242–1253. [PubMed: 25375928]
33. Peng R, Zhang PF, Yang X, Huang XY, et al. Overexpression of RNF38 Facilitates TGF- β Signaling by Ubiquitinating and Degrading AHNK in Hepatocellular Carcinoma. *J Exp Clin Cancer Res*. 2019; 38.1: 113. [PubMed: 30836988]
34. Palomino DC and Marti LC. Chemokines and Immunity. *Einstein (Sao Paulo)*. 2015; 13.3: 469–473. [PubMed: 26466066]
35. Jovov B, Araujo-Perez F, Sigel CS, Stratford JK, et al. Differential Gene Expression between African American and European American Colorectal Cancer Patients. *PLoS One*. 2012; 7.1: e30168. [PubMed: 22276153]

36. Boland PM and Ma WW. Immunotherapy for Colorectal Cancer. *Cancers (Basel)*. 2017; 9.5: 50.
37. Franke AJ, Skelton WP, Starr JS, Parekh H, et al. Immunotherapy for Colorectal Cancer: A Review of Current and Novel Therapeutic Approaches. *J Natl Cancer Inst*. 2019; 111.11: 1131–1141. [PubMed: 31322663]
38. Spratt DE, Chan T, Waldron L, Speers C, et al. Racial/Ethnic Disparities in Genomic Sequencing. *JAMA Oncol*. 2016; 2.8: 1070–1074. [PubMed: 27366979]
39. Kosti I, Jain N, Aran D, Butte AJ, et al. Cross-Tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues. *Sci Rep*. 2016; 6: 24799. DOI: 10.1038/srep24799. [PubMed: 27142790]
40. Mersha TB and Abebe T. Self-Reported Race/Ethnicity in the Age of Genomic Research: Its Potential Impact on Understanding Health Disparities. *Hum Genomics*. 2015; 9.1. DOI: 10.1186/s40246-014-0023-x.

Figure 1:

IPA Enrichment Analysis of Differences in Canonical Signaling Pathways, BA compared to NHWA. Depicted are the canonical signaling pathways significantly enriched with differential expression between BA and NHWA colon cancer specimens. Each bar represents a difference in expression of a canonical signaling pathway. The corresponding p-value for the magnitude of the difference in expression (activation, suppression, or unknown) is represented by the width of the bar. Green bars indicate IPA-predicted suppression of the canonical pathway activity in BA colon cancer specimens relative to NHWA colon cancer specimens. White bars indicate unknown suppression or activation of the canonical pathway in BA colon cancer specimens relative to NHWA colon cancer specimens. Each circle indicates the number of differentially expressed genes in each canonical signaling pathway.

Figure 2. IPA Enrichment Analysis of Differences in Biological Processes, BA compared to NHWA. Depicted are the biological processes significantly enriched with differential expression between BA and NHWA colon cancer specimens. Each bar represents a difference in expression of a biological process. The corresponding p-value for the magnitude of the difference in expression (activation, suppression, or unknown) is represented by the width of the bar. Red bars indicate IPA-predicted activation of the biological process in BA colon cancer specimens relative to NHWA colon cancer specimens. Green bars indicate IPA-predicted suppression of the biological process in BA colon cancer specimens relative to NHWA colon cancer specimens. White bars indicate unknown suppression or activation of the biological process in BA colon cancer specimens relative to NHWA colon cancer specimens. Each circle indicates the number of differentially expressed genes in each biological process.

Table 1:

Patient Demographics

Specimen #	Gender	Age at Time of Surgery, yrs	Colon Cancer Stage	Location of Primary Tumor	Tumor Grade	Lymphovascular Invasion
BA Colon Cancer Specimens						
5	Female	79	IIA	Ascending	Low	No
10	Female	68	IIIB	Descending	Low	Yes
12	Male	60	I	Ascending	Low	No
14	Female	76	IIA	Ascending	Low	No
16	Male	51	IIA	Descending	Low	Yes
18	Female	64	IIA	Ascending	Low	No
21	Female	53	IIB	Ascending	Low	No
24	Male	74	IIIB	Ascending	Low	Yes
NHLA Colon Cancer Specimens						
6	Female	52	I	Descending	Low	No
8	Female	62	IIIB	Ascending	Low	Yes
9	Male	52	IIIB	Ascending	Low	Yes
11	Male	78	IIIB	Descending	Low	Yes
13	Female	60	IIB	Descending	Low	Yes
15	Male	76	IIA	Descending	Low	No
17	Female	45	I	Descending	Low	No
20	Female	76	IIIB	Ascending	High	Yes
22	Male	71	IIIB	Ascending	Low	Yes

Table 2:

Top 20 Up-Regulated Genes, BA Compared to NHPA

Gene Symbol	Gene Name	Chromosome Location	Log ₂ (BA/NHPA)	q-value
SOX30	SRY-Box Transcription Factor 30	chr5:157050822–157107162	13.7827	0.00064321
TRPM3	Transient Receptor Potential Cation Channel Subfamily M Member 3	chr9:73149368–73736514	10.7593	0.00018014
REXO1	RNA Exonuclease 1 Homolog	chr19:1753518–1876164	6.35986	0.00018014
MUC6	Mucin 6	chr11:1012823–1036706	5.69516	0.00018014
CPSF1	Cleavage And Polyadenylation Specific Factor 1	chr8:145597703–145634733	4.11069	0.00018014
ANKFN1	Ankyrin Repeat And Fibronectin Type III Domain Containing 1	chr17:54188303–54572793	6.07007	0.00406045
TKTL1	Transketolase Like 1	chrX:153524026–153558713	5.38324	0.00219301
NEK10	NIMA Related Kinase 10	chr3:27151494–27411533	3.79868	0.00018014
IGF2	Insulin Like Growth Factor 2	chr11:2150341–2182439	3.93341	0.00033858
GRPR	Gastrin Releasing Peptide Receptor	chrX:16041862–16202231	3.29795	0.00018014
DEFA6	Defensin Alpha 6	chr8:6782215–6783598	3.2816	0.00018014
GREB1L	GREB1 Like Retinoic Acid Receptor Coactivator	chr18:18822202–19105723	3.70072	0.00049117
SGPP2	Sphingosine-1-Phosphate Phosphatase 2	chr2:223288684–223436036	3.14134	0.00018014
CCL25	C-C Motif Chemokine Ligand 25	chr19:8117252–8127547	4.67392	0.00374859
ASMTL	Acetylserotonin O-Methyltransferase Like	chrX:1516567–1572655	4.9107	0.00670374
CASC18	Cancer Susceptibility 18	chr12:106069631–106179173	7.84356	0.0440943
OLFM3	Olfactomedin 3	chr1:102268122–102462790	5.26474	0.0107215
PAGE1	PAGE Family Member 1	chrX:49452053–49460596	6.73934	0.0369055
EYA1	EYA Transcriptional Coactivator And Phosphatase 1	chr8:72007541–72274514	3.68486	0.00250468
UGT2B4	UDP Glucuronosyltransferase Family 2 Member B4	chr4:70345882–70361626	4.01003	0.00421581

BA = Black American; NHPA = Non-Hispanic White American; Chr = Chromosome

Table 3:

Top 20 Down-Regulated Genes, BA Compared to NHA

Gene Symbol	Gene Name	Chromosome Location	Log ₂ (BA/NHA)	q-value
RNF38	Ring Finger Protein 38	chr9:36316761–36487651	–11.4695	0.00018014
GVINP1	GTPase, Very Large Interferon Inducible Pseudogene 1	chr11:6727194–6979583	–8.28241	0.00018014
TMEM72-AS1	TMEM72 Antisense RNA 1	chr10:45306471–45455137	–8.07589	0.029061
LOC102723769	Uncharacterized LOC102723769	chr1:143119057–143236267	–7.72515	0.00018014
ZBED3	Zinc Finger BED-Type Containing 3	chr5:76365675–76444176	–6.20325	0.00018014
AQP5	Aquaporin 5	chr12:50344523–50359866	–5.8446	0.0008001
MIR940	MicroRNA 940	chr16:2321747–2322890	–5.73655	0.00871762
GARNL3	GTPase Activating Rap/Ran GAP Domain Like 3	chr9:129986752–130156198	–5.31645	0.00018014
SERPINB3	Serpin Family B Member 3	chr18:61322430–61329197	–5.12972	0.00421581
METTL12	Methyltransferase-Like Protein 12	chr11:62414319–62457371	–4.93269	0.00018014
PER1	Period Circadian Regulator 1	chr17:8043786–8059682	–4.67949	0.00018014
ZIC1	Zinc Family Member 1	chr3:147127180–147134506	–4.63001	0.00204412
PRSS21	Serine Protease 21	chr16:2867163–2871723	–4.58465	0.00018014
SNHG11	Small Nucleolar RNA Host Gene 11	chr20:37075296–37079583	–4.43045	0.00018014
LOC154761	Family With Sequence Similarity 115, Member C	chr7:143257418–143599278	–4.03778	0.00018014
DGAT1	Diacylglycerol O-Acyltransferase 1	chr8:145515269–145550582	–3.94599	0.00018014
DEPDC4	DEP Domain Containing 4	chr12:100636313–100660857	–3.83914	0.00033858
DES	Desmin	chr2:220283098–220291461	–3.71151	0.00018014
UBA7	Ubiquitin Like Modifier Activating Enzyme 7	chr3:49842637–49851391	–3.69744	0.00018014
EVC	EvC Ciliary Complex Subunit 1	chr4:5712888–5894810	–3.23094	0.00018014

BA = Black American; NHA = Non-Hispanic White American; Chr = Chromosome

Table 4:

SOX30, TRPM3 and RNF38 Gene Expression in Different Cancer Tissues, BA Compared to NHWA,
Extracted from The Cancer Genome Atlas

Gene Symbol	Cancer	Log ₂ (BA/NHWA)	N for BA	N for NHWA	Adjusted p-value
SOX30	BLCA	1.1841	22	329	0.11202
SOX30	HNSC	-0.3584	46	425	0.376901
SOX30	KIRK	0.9527	55	466	0.034657
SOX30	LUAD	0.2226	53	404	0.675349
SOX30	LUSC	-0.0715	29	349	0.935206
SOX30	UCEC	0.7472	107	377	0.000654
TRPM3	BLCA	-2.0786	22	329	0.069567
TRPM3	HNSC	0.1936	46	425	0.741499
TRPM3	KIRK	0.1132	55	466	0.634292
TRPM3	LUAD	0.2751	53	404	0.764888
TRPM3	LUSC	-0.2757	29	349	0.557711
TRPM3	UCEC	0.2984	107	377	0.387
RNF38	BLCA	0.1832	22	329	0.425175
RNF38	HNSC	0.0781	46	425	0.66287
RNF38	KIRK	0.0272	55	466	0.834408
RNF38	LUAD	-0.1743	53	404	0.090479
RNF38	LUSC	-0.3244	29	349	0.013948
RNF38	UCEC	0.1074	107	377	0.188246

BA = Black American; NHWA = Non-Hispanic White American; N = Number; BLCA = Bladder Carcinoma; HNSC = Head-Neck Squamous Cell Carcinoma; KIRC = Kidney Renal Clear Cell Carcinoma; LUAD = Lung Adenocarcinoma; LUSC = Lung Squamous Cell Carcinoma; UCEC = Uterine Corpus Endometrial Carcinoma.

Bold, italicized entries are statistically differentially expressed between BA and NHWA.