NEOPLASIA
www.neoplasia.com

Volume 20 Number 3 March 2018 pp. 289–294 289

Single Nucleotide Polymorphism Facilitated Down-Regulation of the Cohesin Stromal Antigen-1: Implications for Colorectal Cancer Racial Disparities

Somenath Datta*, Richard M. Sherva†, Mart De La Cruz*, Michelle T. Long*, Priya Roy*, Vadim Backman‡, Sanjib Chowdhury*,1 and Hemant K. Roy*,1

*Department of Medicine, Section of Gastroenterology, Boston University Medical Center, Boston, MA 02118, USA; †Department of Medicine, Section of Biomedical Genetics, Boston University Medical Center, Boston, MA 02118, USA; ‡Department of Biomedical Engineering, Northwestern University, Evanston, Illinois, USA

Abstract

The biological underpinnings for racial disparities in colorectal cancer (CRC) incidence remain to be elucidated. We have previously reported that the cohesin SA-1 down-regulation is an early event in colon carcinogenesis which is dramatically accentuated in African-Americans. In order to investigate the mechanism, we evaluated single nucleotide polymorphisms (SNPs) for association with SA-1-related outcomes followed by gene editing of candidate SNP. We observed that rs34149860 SNP was significantly associated with a lower colonic mucosal SA-1 expression and evaluation of public databases showed striking racial discordance. Given that the predicted SNP would alter miR-29b binding site, we used CRISPR knock-in in CRC cells and demonstrated that the SNP but not wild-type had profound alterations in SA-1 expression with miR-29b inhibitor. This is the first demonstration of high-order chromatin regulators as a modulator of racial differences, risk alteration with SNPs and finally specific modulation by microRNAs.

Neoplasia (2018) 20, 289–294

Introduction

Colorectal cancer (CRC) remains the second leading causes of cancer mortality in the United States and Western countries despite effective modalities for screening fecal tests and colonoscopy [1]. Barriers to implementation include compliance, resource utilization and complication/discomfort. This is juxtaposed by the relatively low rate (~7%) of advanced adenomas, the immediate precursor of CRC, meaning that the vast majority of colonoscopies do not have cancer preventive implications. This underscores the need of more effective risk prediction strategies to personalize CRC screening strategies [2]. Approximately 30% of CRC risk is considered genetic making inherited factors a logical target for risk stratification efforts [3].

Genetic predilection for CRC is epitomized by high penetrance but relatively infrequent syndromes such as familial adenomatous polyposis and Lynch syndrome. More commonly, risk alleles are due to single nucleotide polymorphisms (SNPs) in genes important in colon carcinogenesis. Efforts to harness this have been through genome-wide association studies (GWAS) with a plethora of report albeit most with low risk factors. Furthermore, variation by population has also been reported in GWAS.

Since many SNPs have a racial predilection, this provides a potential mechanism for racial disparities in CRC. Strikingly the relative incidence and mortality rates of CRC were ~25% and ~50% higher, respectively, in Blacks compare to Whites [1,4]. While socio-economic factors are undoubtedly contributory, there are also marked biological differences between CRCs in Blacks and Whites [5]. These might be genetically related with studies showing SNPs in p53 [6], TGFβ-related genes (BMP4, GREM1, CDH1, SMAD7 and RPHN2) [7] and mannose-binding lectin 2 (MBL2) gene [8].

Address all correspondence to: Sanjib Chowdhury, PhD, MS, Research Assistant Professor of Medicine, Boston University School of Medicine or Hemant K Roy MD, Franz Ingelfinger Professor of Medicine, Boston University School of Medicine, Chief, Section of Gastroenterology, Boston Medical Center, 650 Albany Street, Boston, MA, 02118.
E-mails: sanjib@bu.edu, hkroy@bu.edu
[1] Contributed equally.
Received 9 October 2017; Revised 21 December 2017; Accepted 3 January 2018

© 2018 The Authors. Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1476-5586
https://doi.org/10.1016/j.neo.2018.01.003

Given the myriad of differentially expressed genes dysregulated, high order chromatin remodellors as a modality for these pleotropic genomic effects [9–12]. The cohesins, including SA-1 (also known as *STAG1* and *SCC3*), are a recently discovered high order chromatin remodellors that are involved in chromatin looping (bringing enhancers/promoters in proximity to coding regions to foster transcription) and thus facilitate long-range nuclear interactions and gene expression [13]. While mutations in this family of proteins are rare in CRC, we have recently reported that SA-1 was significantly down-regulated in the endoscopically normal rectal mucosa of patients harboring adenomas elsewhere in their colon (field carcinogenesis) presumably through epigenetic mechanisms [14]. Moreover, SA-1 loss triggered critical pathways in carcinogenesis [15–17]. Importantly, we noted that rectal mucosal SA-1 was dramatically reduced in Blacks when compared to Whites providing a potential mechanism for racial differences in CRC rates. Since SNPs may segregate racially, we aimed to test the hypothesis that SNPs may relate to decrease in SA-1 expression selectively by race. We combined *in silico* work, human colon, and cell culture (CRISPR knock-in) work to demonstrate that an exonic polymorphism rs34149860 (selected based on p-value and racial segregation) had a racial predilection and was associated with decreased SA-1 expression. Mechanistically, we discovered that the rs34149860 impacted miR-29b binding leading to loss of SA-1 mRNA expression.

Materials and Methods

Human Studies

We recruited 81 patients undergoing screening/surveillance colonoscopy and obtained two biopsies from the endoscopically normal rectal mucosa. We also evaluated CRC samples (n = 14) from our repository.

SNP Selection

SA-1 SNPs (n = 27) were selected for genotyping based on: (1) eQTL potential for SA-1 expression in any tissue based on the GTEx database (http://www.gtexportal.org/home/), (2) missense mutations within SA-1 identified in the SnpEffdatabases (http://snpeff.sourceforge.net/), (3) SNPs capturing the haplotype diversity in SA-1.

SNP Array

SNP genotyping was performed using 7900HT Fast Real-Time PCR System (Life Technologies, Grand Island, NY) and analyzed using SDS software version 2.3.

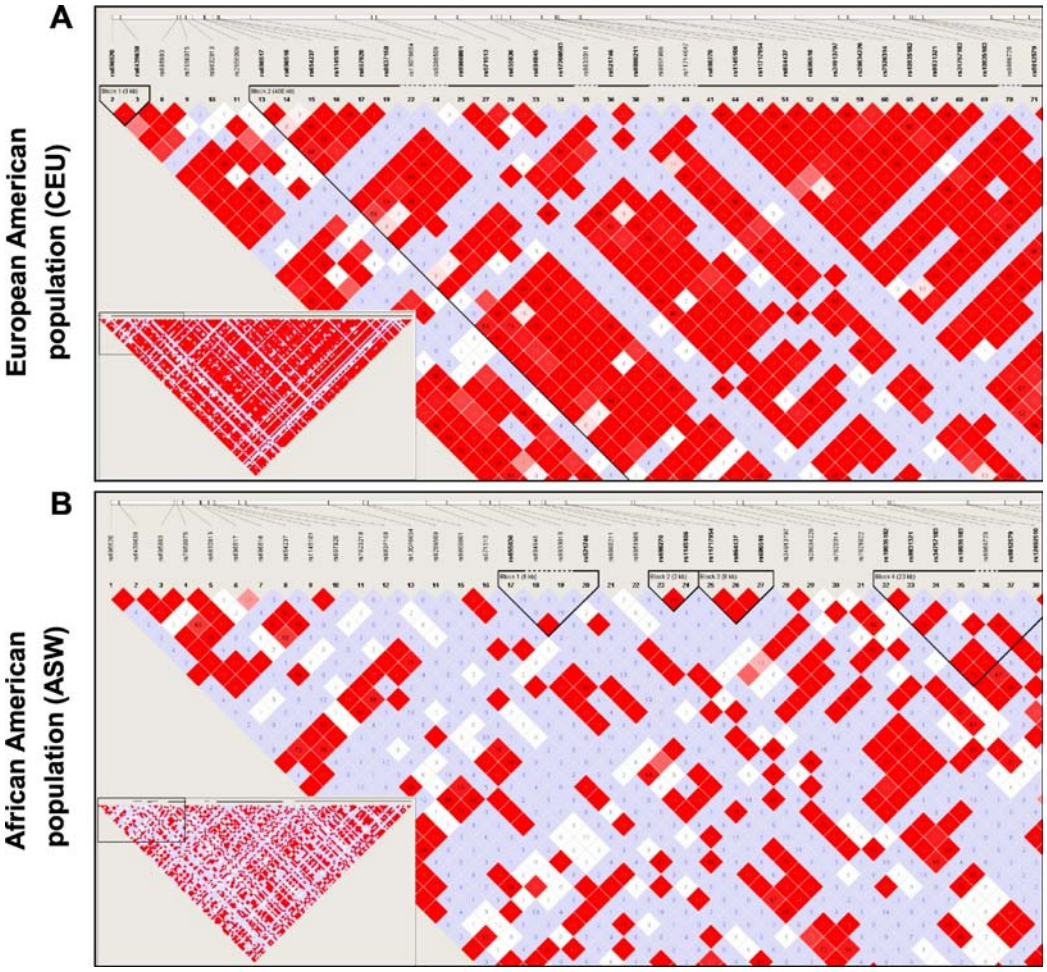

Figure 1. Cohesin SA-1 (*STAG1*) gene mapping for the (A) CEU and (B) ASW populations formed by genotyped SNPs. Abbreviations: CEU – European American (White) population represented by Utah residents with Northern and Western European ancestry. ASW – African American (Black) population represented by the African ancestry in Southwest USA.

Table 1. Effect of SA1 Polymorphisms on SA-1 Expression in Total Patient Cohort

Association of SNPs With SA-1 Expression						
SNP	SA1 Deregulation	*P*-Value	% AFR	% EUR	Type of Variant	Selection Method
rs34149860	161% down	7.992e-05	31	0	Missense	Missense
rs1070232	125% down	0.03096	30	26	Intron	tag SNP
rs13071220	143% down	0.00073	7	20	Intron	tag SNP
rs1681818	197% down	0.0011	31	26	Non coding transcript exon variant	tag SNP
rs17364492	115% down	0.01076	14	20	Intron	tag SNP
rs2655008	159% down	0.00012	33	27	Intron	tag SNP
rs34992220	199% down	0.00018	44	67	Intron	tag SNP
rs4678433	197% down	0.0023	34	29	Intron	tag SNP
rs4678437	128% down	0.048	48	31	Intron	tag SNP
rs6773604	150% down	0.00016	40	10	Intron	eQTL
rs696517	141% down	0.0003	27	20	Non coding transcript exon variant	tag SNP
rs7646127	109% down	0.03778	32	0	Intron	tag SNP
rs9861150	133% down	0.0184	14	22	Intron	tag SNP

Abbreviations: *AFR*, African; *EUR*, European.
The racial prevalence estimate (% AFR and % EUR) are obtained from well-annotated databases (dbSNP and Ensembl).

CRISPR-KI Method

CRISPR gene knock-in technique was used to change the SNP, rs34149860 in the SA-1 (*STAG1*) genomic DNA in HCT116 and RKO cells. Gene expression performed using Step-One-Plus RT Thermo cycler (Life Technologies).

DNA Sequencing

Genotyping of CRC cells for SNP rs34149860 and validation of CRISPR knock-in was done by PCR amplification and purification followed by Sanger DNA sequencing at Genewiz (Cambridge, MA) using the following primers: rs34149860.Fwd: CAGAGCCA CACTCTTCACAAGT; rs34149860.Rev: ACCTGGTGAGACA CCGTACATT.

Statistical Methods

Logistic and linear regression models (PLINK) were used for the analysis to test for associations between SA-1 SNPs, SA-1 expression, and CRC risk. Interaction tests were used to determine whether the effect of these SNPs differed between Blacks and Whites. To assess the significance of the observed *P* values and correct for multiple testing, we created 1000 replicates of the dataset by randomly permuting genotypes and phenotypes.

Results

In Silico Analysis Linkage Disequilibrium and Haplotype Diversity in Cohesin SA-1 Between Blacks and Whites

Linkage disequilibrium (LD) patterns have been well-characterized in population-genetics studies as well as mapping disease-causing genes by comparing allele frequencies between case and control groups. The LD across the different regions of the world varies considerably due to biological and random factors and is partly attributed to age of population and the number of founders. Based on our recent report demonstrating significant loss of SA-1 gene expression (both baseline and in cancer) in Black population, we investigated the LD distribution of SA-1 gene between Blacks and Whites within the general population using public databases.

We downloaded genome-wide SNP genotyping datasets using HapMap 3 (International HapMap Project, Phase 3 samples) Genome Browser release #28 and used the Haploview 4.2 software tool (Broad Institute/MIT) to compute linkage disequilibrium (LD) and population haplotype patterns for Blacks represented by the African ancestry in Southwest USA (ASW; n = 90) cohort and Whites represented by Utah residents with Northern and Western European ancestry from the CEPH collection (CEU; n = 180). As shown in Figure 1*A* and *B*, we observed striking heterogeneity in the LD (r^2) between the Black and White populations within the genomic region of SA-1.

Effect of SA-1 SNPs on SA-1 Gene Expression With Race-Specific Implication

Based on the dramatic difference in the LD of SA-1 gene, we next investigated whether SA-1 polymorphisms are associated with the loss of SA-1 gene expression and risk of CRC. The SA-1 gene is a member of the SCC3 family and expressed in the nucleus. It is located on Chromosome 3 (3q22.3). The gene contains 34 exons (33 coding exons) with a transcript length of 6055bps and translational length of 1258 amino acid residues. We observed 23,484 SNPs in human SA-1. Out of these SNPs, 236 were missense mutations, 22,789 were introns, 146 were 3'UTR and 89 were 5' UTR. To identify a candidate SA-1 variant that is associated with decreased gene expression and risk of CRC, we selected 27 SNPs based on eQTL potential, impact on protein function, and haplotype tagging. Using Taqman SNP genotyping array we have identified 13 SNPs which were found significantly associated with SA-1 down-regulation (without race adjustment) (Table 1). We focused on SNP rs34149860 which encoded a missense alteration in SA-1 (allele change – CA**G** = CA**C**; residual change – **Q** [Gln] = **H** [His] because it was observed to have marked racial prevalence differences (31% of black and 0% of whites). Furthermore, when we evaluated our 14 CRC patients, we also observed that rs34149860 was associated with CRC risk (OR: 1.96; 95% CI: 0.70 ± 5.453; $P = 0.2$). These data show for the first time that SA-1 SNPs have racial predilection providing a biological basis into the preferential loss of the SA-1 tumor suppressor gene in African Americans and hence higher incidence of CRC.

In-Vitro Engineering of rs34149860 SNP Through CRISPR Knock-In

We genotyped a panel of CRC cells for presence of SNP rs34149860. RKO cells which are heterozygous for the SNP

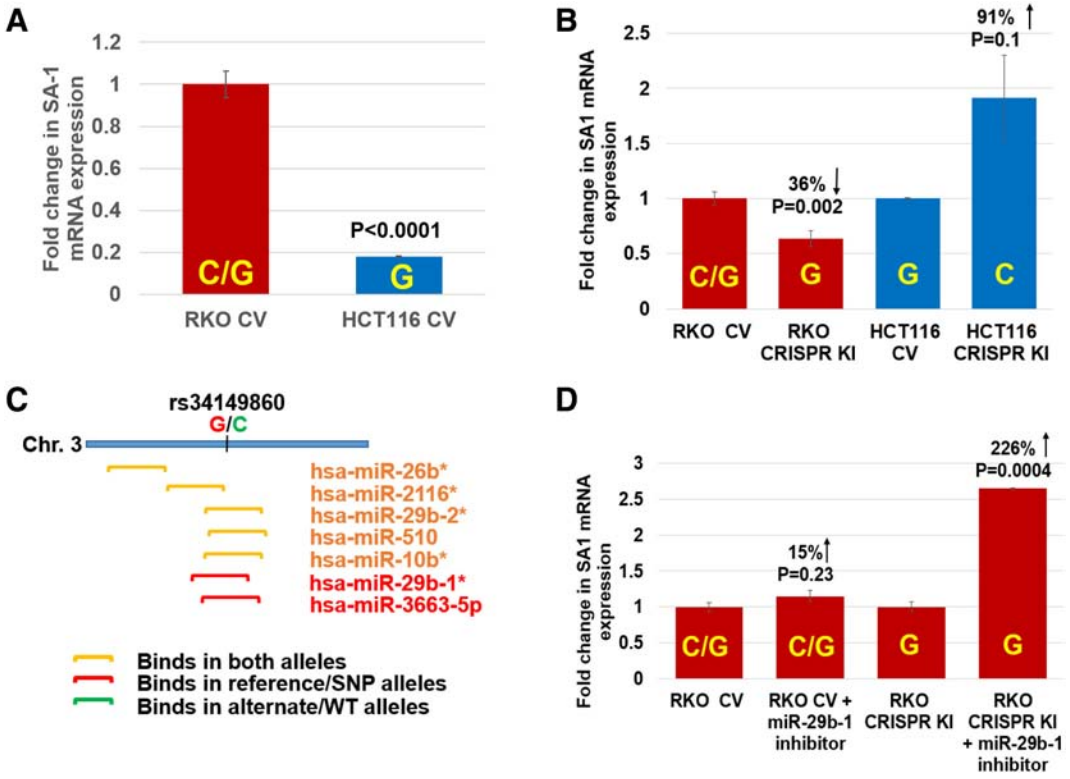

Figure 2. Impact of rs34149860 on SA-1 Expression. (A) SA-1 mRNA expression in RKO and HCT116 cells. (B) Conversion of C (alternate allele) to G (reference allele) with CRISPR leads to decreased SA-1 Expression. (C) miR-29b is predicted to have differential binding of SA-1. D. miR-29b inhibitor markedly effected SA-1 expression in reference allele (G) but not in alternate allele (C).

rs34149860 (CG, one alternate allele—C and one reference allele G) showed higher SA-1 expression compare to HCT116 cells which are homozygous for the SNP (GG) (Figure 2*A*). We successfully designed CRISPR knock-in to create the reciprocal allele alterations within RKO and HCT116 cells. In RKO cells, changing to a homozygous variant (CG to GG) alleles decreased SA-1 by 36% ($P = .002$) whereas in HCT116 cells changing from reference allele variant to alternate allele (GG to CC) increased expression by 91% (Figure 2*B*). These results suggest a significant contribution of rs34149860 SNP in the regulation of SA-1 expression.

miRNA Mediated Down-Regulation of SA-1 Targeting SNPs

Our discovery that rs34149860 SNP is associated with decreased SA-1 expression is intriguing. However, it is not intuitive to believe a missense alteration going from glutamine➔histidine will result in decreased SA-1 mRNA expression as noted. Since it does not involve promoter or enhancer regions but rather exons, it is unlikely to be methylation. We posited that SNP rs34149860 might play a role in miRNA (miR) binding and mRNA destabilization. Presence of SNPs can disrupt normal miR binding and/or introduce new binding sites which might be associated with disease pathogenesis. This association emphasizes the importance of detecting miR targets and predicting the possible effects of SNPs on binding sites. Dysregulation of miRs are common in colon carcinogenesis [18]. Our group has noted that miRs are dysregulated in field carcinogenesis [19]. We used RegRNA (A Regulatory RNA Motifs and Elements Finder) to identify the predicted miRNA binding to the rs34149860 region (±25bp around SNP transcript) that would be disrupted by variant. Our analysis on the possible effects of SNPs on miR binding sites [20] showed binding of miR-29b-1* as the most promising candidate (Figure 2*C*). In order to get empiric data, we used RKO cells (baseline heterozygous) and engineered variant status (homozygous) with CRISPR knock-in. We noted that the induction of SA-1 expression with miR-29b-1* inhibitor showed a minimal (15%. $P = .23$) increase in the control cells whereas the same concentration of inhibitor showed a dramatic 226% increase ($P = .004$) in the variant (Figure 2*D*). To determine the clinical significance, we validated that miR-29b-1* was expressed in rectal mucosa of ~30 randomly selected patients (data not shown). These emerging findings are highly novel and provide novel mechanistic insights into CRC racial disparities.

Discussion

In this proof of concept study, we took a candidate approach to identify a SA-1 SNP that was associated with decreased SA-1 expression in the colonic mucosa. We report herein, that SNP rs34149860 enabled miR-mediated degradation of colonic mucosal cohesin SA-1 expression which we previously reported is the hallmark of CRC risk. Importantly, the racial predilection of this SNP provides a compelling potential mechanism for CRC disparities.

High order chromatin has been shown to be a fundamental event in carcinogenesis. We have demonstrated that high order chromatin is altered in the endoscopically-normal rectal mucosa patients harboring colonic neoplasia using partial wave spectroscopic microscopy (PWS) [21]. Members of SWI/SNF family including Arid 1a has been shown to be one of the most commonly mutated genes involved in chromatin-remodeling and gene regulation [22,23]. A new family of high order chromatin remodelers is the cohesins. These exist as a family (including Smc1, Smc3, Rad21/Scc1 and SA-1 or SA2) that

can play a critical role in sister chromatid segregation (cohesion) and more relevant to CRC initiation is change in high order chromatin allowing genes to be transcribed via topologically-associated domains and chromatin loops [13,15,24,25]. A case in point has been a seminal report on glioblastoma that human isocitrate dehydrogenase 1 (IDH-1) mutant gliomas exhibit hypermethylation at cohesin and CCCTC-binding factor (CTCF)-binding sites, compromising binding of this methylation-sensitive insulator protein [16]. Reduced CTCF binding is associated with loss of insulation between topological domains and aberrant gene activation through allowing interaction of a constitutive enhancer element with PDGFRA. These cohesins are turned off with factors such as WAPL providing a mechanism to restrict chromatin looping functions through triggering cohesin release from chromatin [25,26]. The ability of cohesins to impacting promoter-enhancer function is potentially apropos given the emerging data on the role of enhancers in colon carcinogenesis [9]. In this regard, Cohen et al. showed that approximately half of the GWAS CRC risk loci co-localized with recurrently activated enhancers. Moreover, several of these enhancers are occupied by cohesin complex members [9]. Additionally, it has been noted that enhancer activation via Arid1a loss is sufficient to induce intestinal tumorigenesis thus emphasizing the potential biological implications [10–12].

While somatic mutations in cohesins are well established, germline mutations or SNPs as a risk factor has not been described. To our knowledge this is the first report of SNPs in chromatin remodellors associated with colon cancer risk and thus may provide important insights. Furthermore, the racial predilection may explain, at least partly, the striking differences in CRC rates for Blacks compared to Whites. The patient database was small and thus the odds ratio for CRC was not significant but the decreased colonic SA-1 expression is compelling given the unequivocal data that SA-1 down-regulation is an earliest event in colon carcinogenesis [27].

MicroRNAs (miRs) are short ~22-27 nucleotide long non-coding RNAs that bind specific sequences on mRNA and promote degradation through RNA-induced silencing complex (RISC). Dysregulation of miRs is an important epigenetic mediator during colon carcinogenesis and has been exploited as a biomarker (found in blood and stool). We have earlier reported that miRs are dysregulated in field carcinogenesis [19] thus providing a mechanism for SA-1 degradation. In this proof of concept study, we documented the novel interaction between SA-1 and miR which is based partly on the critical hypothesis that miRs may act selectively on SNPs or certain miRs may be selectively altered in the colon by race. In addition, miR-29b has been shown to be an important CRC biomarker supporting its relevance [28].

There are numerous strengths of this study. To our knowledge, this is the first to link SNPs in cohesins in particular and high order chromatin modulators in general to CRC risk. This is particularly important given the emerging knowledge regarding the importance of these. The comparison to actual mRNA in in the colonic mucosa is novel and powerful in a manner to go beyond the typical association studies. The elucidation of a mechanism through which SNP leads to decreased RNA expression provides biological plausibility and insights.

There are several limitations that need to be acknowledged. The SNP selection was based on candidate approach and not comprehensive with other SNPs may also be important. It bears reiteration that our goal was simply to demonstrate proof of principle that SNPs can impact SA-1 expression and CRC risk. Therefore, we deferred standard biomarker approaches (independent validation set, adjustment for multiple comparators etc.) for future studies. There might be other epigenetic modulations including methylation that may also contribute to the racial differences. The link between SA-1 and CRC is indirect based on the compelling association between mucosal SA-1 and CRC risk [27]. Our pilot CRC data (n = 14) was suggestive (~2-fold risk) but failed to achieve statistical significance. Furthermore, given that self-reported race [29] accuracy has been recently impugned and we were unable to genotype our patients, we resorted upon well-annotated databases for racial prevalence estimate since genotyping our patients was not feasible.

In conclusion, we provide the first report on cohesin SA-1 SNP and colonic mucosal gene expression as a surrogate for CRC risk. The mechanistic insights into the role of miRNA mediated SA-1 degradation provide novel insights. While corroboration in large scale is needed for rs34149860 and other SA-1 SNPs, this has potential for personalization of screening targeting racial differences.

Acknowledgements

We thank Ms. Irene Simkin, Project Manager, Molecular Genetics Core at Boston University School of Medicine for performing the SNP array and Ms. Beth Parker for excellent support in manuscript preparation.

Grant Support

NIH grants – R03CA195143, R01CA165309, R01CA156186, R01CA200064, R01CA183101, R33CA225323-01.

The Role of the Study Sponsor

The National Institutes of Health had no role in the study.

Conflict of Interest Statement

Drs. Roy and Backman are co-founders and shareholders in Nanocytomics LLC. All other authors do not have any competing interests.

References

[1] Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RG, Barzi A, and Jemal A (2017). Colorectal cancer statistics. *CA Cancer J Clin* **67**, 177–193.

[2] Song M and Giovannucci E (2016). Preventable Incidence and Mortality of Carcinoma Associated With Lifestyle Factors Among White Adults in the United States. *JAMA Oncol* **2**(9), 11.

[3] Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, Graff RE, Holst K, Moller S, and Unger RH, et al (2016). Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *JAMA* **315**(1), 68–76.

[4] Siegel RL, Miller KD, and Jemal A (2017). Cancer Statistics, 2017. *CA Cancer J Clin* **67**(1), 7–30.

[5] Guda K, Veigl ML, Varadan V, Nosrati A, Ravi L, Lutterbaugh J, Lawrence E, Lewis S, Willson JK, and Lowe JB, et al (2015). Novel recurrently mutated genes in African American colon cancers. *Proc Natl Acad Sci U S A* **112**(4), 1149–1154.

[6] Katkoori VR, Jia X, Shanmugam C, Wan W, Meleth S, Bumpers H, Grizzle WE, and Manne U (2009). Prognostic significance of p53 codon 72 polymorphism differs with race in colorectal adenocarcinoma. *Clin Cancer Res* **15**(7), 2406–2416.

[7] Kupfer SS, Skol AD, Hong E, Ludvik A, Kittles RA, Keku TO, Sandler RS, and Ellis NA (2014). Shared and independent colorectal cancer risk alleles in TGFbeta-related genes in African and European Americans. *Carcinogenesis* **35**(9), 2025–2030.

[8] Zanetti KA, Haznadar M, Welsh JA, Robles AI, Ryan BM, McClary AC, Bowman ED, Goodman JE, Bernig T, and Chanock SJ, et al (2012). 3'-UTR and functional secretor haplotypes in mannose-binding lectin 2 are associated

with increased colon cancer risk in African Americans. *Cancer Res* **72**(6), 1467–1477.

[9] Cohen AJ, Saiakhova A, Corradin O, Luppino JM, Lovrenert K, Bartels CF, Morrow JJ, Mack SC, Dhillon G, and Beard L, et al (2017). Hotspots of aberrant enhancer activity punctuate the colorectal cancer epigenome. *Nat Commun* **8**, 14400.

[10] Mathur R, Alver BH, San Roman AK, Wilson BG, Wang X, Agoston AT, Park PJ, Shivdasani RA, and Roberts CW (2017). ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat Genet* **49**(2), 296–302.

[11] Lakshminarasimhan R, Andreu-Vieyra C, Lawrenson K, Duymich CE, Gayther SA, Liang G, and Jones PA (2017). Down-regulation of ARID1A is sufficient to initiate neoplastic transformation along with epigenetic reprogramming in non-tumorigenic endometriotic cells. *Cancer Lett* **401**, 11–19.

[12] *ARID1A Deficiency Promotes Colorectal Cancer via Enhancer Dysregulation* (2017).*Cancer Discov* **7**(2), OF7.

[13] Ohlsson R (2010). Gene expression: The coherent Mediator. *Nature* **467**(7314), 406–407.

[14] Wali RK, Momi N, Dela Cruz M, Calderwood AH, Stypula-Cyrus Y, Almassalha L, Chhaparia A, Weber CR, Radosevich A, and Tiwari AK, et al (2016). Higher Order Chromatin Modulator Cohesin SA1 Is an Early Biomarker for Colon Carcinogenesis: Race-Specific Implications. *Cancer Prev Res (Phila)* **9**(11), 844–854.

[15] Losada A (2014). Cohesin in cancer: chromosome segregation and beyond. *Nat Rev Cancer* **14**(6), 389–393.

[16] Grimmer MR and Costello JF (2016). Cancer: Oncogene brought into the loop. *Nature* **529**(7584), 34–35.

[17] Anazawa Y, Arakawa H, Nakagawa H, and Nakamura Y (2004). Identification of STAG1 as a key mediator of a p53-dependent apoptotic pathway. *Oncogene* **23**(46), 7621–7627.

[18] Kara M, Yumrutas O, Ozcan O, Celik OI, Bozgeyik E, Bozgeyik I, and Tasdemir S (2015). Differential expressions of cancer-associated genes and their regulatory miRNAs in colorectal carcinoma. *Gene* **567**(1), 81–86.

[19] Kunte DP, DelaCruz M, Wali RK, Menon A, Du HY, Stypula Y, Patel A, Backman V, and Roy HK (2012). Dysregulation of MicroRNAs in Colonic Field Carcinogenesis: Implications for Screening. *PLoS One* **7**(9).

[20] Deveci M, Catalyurek UV, and Toland AE (2014). mrSNP: software to detect SNP effects on microRNA binding. *BMC Bioinform* **15**, 73.

[21] Damania D, Roy HK, Subramanian H, Weinberg DS, Rex DK, Goldberg MJ, Muldoon J, Cherkezyan L, Zhu Y, and Bianchi LK, et al (2012). Nanocytology of rectal colonocytes to assess risk of colon cancer based on field cancerization. *Cancer Res* **72**(11), 2720–2727.

[22] Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, and Wyczalkowski MA, et al (2013). Mutational landscape and significance across 12 major cancer types. *Nature* **502**(7471), 333–339.

[23] *Comprehensive molecular characterization of human colon and rectal cancer* (2012).*Nature* **487**(7407), 330–337.

[24] Busslinger GA, Stocsits RR, van der Lelij P, Axelsson E, Tedeschi A, Galjart N, and Peters JM (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**(7651), 503–507.

[25] Haarhuis JHI, van der Weide RH, Blomen VA, Yanez-Cuna JO, Amendola M, van Ruiten MS, Krijger PHL, Teunissen H, Mederna RH, and van Steensel B, et al (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**(4), 693–707.e14.

[26] Ouyang Z, Zheng G, Song J, Borek DM, Otwinowski Z, Brautigam CA, Tomchick DR, Rankin S, and Yu H (2013). Structure of the human cohesin inhibitor Wapl. *Proc Natl Acad Sci U S A* **110**(28), 11355–11360.

[27] Wali RK, Momi N, Dela Cruz M, Calderwood AH, Stypula Y, Almassalha LM, Chhaparia A, Weber CR, Radosevich A, and Tiwari AK, et al (2016). Higher-Order Chromatin Modulator Cohesin SA1 is an Early Biomarker for Colon Carcinogenesis: Race-Specific Implications. *Cancer Prev Res (Phila)* **9**(11), 844–854.

[28] Li L, Guo Y, Chen Y, Wang J, Zhen L, Guo X, Liu J, and Jing C (2016). The Diagnostic Efficacy and Biological Effects of microRNA-29b for Colon Cancer. *Technol Cancer Res Treat* **15**(6), 772–779.

[29] Mersha TB and Abebe T (2015). Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum Genomics* **9**, 1.