African Americans have the highest incidence and mortality rates of colorectal cancer (CRC) of any ethnic group in the United States. Although some of these disparities can be explained by differences in access to care, cancer screening, and other socioeconomic factors, disparities remain after adjustment for these factors. Consequently, an examination of recent advances in the understanding of ethnicity-specific factors, including genetic and environmental factors relating to risk of CRC, the biology of CRC progression, and the changes in screening and mortality, is important for evaluating our progress toward eliminating the disparities. An overarching limitation in this field is the number and sample size of studies performed to characterize the etiological bases of CRC incidence and mortality in African Americans. Despite this limitation, significant differences in etiology are manifest in many studies. These differences need validation, and their impacts on disparities need more detailed investigation. Perhaps most heartening, improvements in CRC screening can be attributed to the smallest difference in CRC incidence between African Americans and whites since the late 1980s. Cancer mortality, however, remains a persistent difference. (Am J Pathol 2018, 188: 291–303; https://doi.org/10.1016/j.ajpath.2017.07.023)

Colorectal cancer (CRC) is the third most common cancer in both men and women in the United States and the second most common cause of cancer-related death. African Americans bear a disproportionate burden, with an incidence of CRC that is >20% higher than in whites and an even larger difference in mortality. In particular, African Americans are more often diagnosed with CRC at an earlier age and with more advanced disease; and African Americans have a greater proportion of CRCs in the proximal colon. Although some of these differences can be explained by access to care, screening, and other socioeconomic factors, a significant portion of the disparity remains after adjustment for these factors. In this review, we examine the recent literature on genetic and environmental risk factors, molecular characteristics of CRC tumors, screening, and cancer-related mortality, and how these factors contribute to our understanding of the colorectal cancer health disparity in African Americans.

Impact of Risk Factors on CRC Incidence

Endoscopic Screening Reduces Cancer Incidence

Data from the Surveillance, Epidemiology, and End Results program of the National Cancer Institute show that CRC incidence has been decreasing in recent years (Figure 1A). The change in CRC incidence is mostly attributed to increased endoscopic examination of the colorectum and the resulting removal of adenomatous and other polyps that are
precancerous lesions. The Surveillance, Epidemiology, and End Results program also reports trends that reflect annual percentage change of rates across certain time segments, which we have included above each corresponding segment. Whites have seen a decrease in CRC incidence since the 1990s. The incidence of CRC in African Americans began to decrease in the early 2000s. The incidence in the United States of CRC in African Americans was 26% higher than whites in 2013. The newest data from the Surveillance, Epidemiology, and End Results program show that African Americans have an 11% higher incidence than whites, the lowest difference since the late 1980s. Although screening has clearly achieved substantial reductions in incidence, less than full compliance, early-onset CRCs, and missed lesions will resist further reductions; therefore, efforts to understand genetic and environmental risk factors continue to be important.

The increased risk of disease and cancer-related death in African Americans has led to changes in the screening recommendations in this population, specifically in lowering the age to begin colonoscopic screening to 45 years. Historically, African Americans have had lower compliance to CRC screening guidelines. Efforts to increase screening have resulted in an increase in compliance, which has been reviewed elsewhere. Increased colonoscopic screening in African Americans has been cited as a major reason for the recent decrease in incidence in this population (Figure 1A).

Continued efforts to increase knowledge of new screening guidelines in African American communities are necessary. In addition, because patient follow-up is lower in African Americans after an abnormality is found, community-based support programs should be developed to improve follow-up rates. More important, access to screening options and affordable care are essential to decrease the burden of CRC faced by African Americans, and these barriers must be addressed by communities across the country.

**Genetic Risk Factors and CRC Incidence**

It is estimated that genetic factors contribute as much as 35% to the overall risk of CRC. Our understanding of genetic risk factors is anchored in mendelian genetics (ie, single-gene defects that are associated with a high risk of CRC development). Mutations in the adenomatous polyposis coli (APC) gene are linked to familial adenomatous polyposis. APC mutations ablate a key factor in the regulation of the WNT signaling pathway (Figure 2). WNTs are a large family of secreted glycoproteins that regulate cell proliferation, differentiation, polarity, and migration. In the adult intestine, WNTs control homeostasis by maintaining stem cell populations in the base of the crypts. In the canonical WNT signaling cascade, WNT stimulation prevents degradation of the transcription factor β-catenin by inhibition of the destruction complex. The destruction complex consists of the APC protein, axis inhibition protein, glycogen synthase kinase 3, and casein kinase 1 isoforms. In the absence of WNT, the complex mediates the phosphorylation of β-catenin, which targets β-catenin to the proteasome.
dominant condition in which individuals inherit a mutation in \textit{APC} from one parent. Mutations in the working \textit{APC} gene copy occur in somatic cells in the colon, producing hundreds to thousands of adenomatous polyps. Biallelic somatic mutations in \textit{APC} are found in approximately 80% of all sporadic CRC cases, making \textit{APC} the predominant gatekeeper gene to the development of CRC.

Recently, biallelic mutations in the base excision repair gene \textit{MUTYH} have been associated with a recessive CRC polyposis syndrome. Base excision repair is a system that removes damaged bases in DNA, particularly oxidized guanine bases, followed by polymerase-associated insertion of the correct nucleotide and ligation. \textit{MUTYH} specifically operates on lesions in which the oxidized guanine base had undergone DNA replication, resulting in an oxidized guanine/adenine mismatch. To prevent mutation, \textit{MUTYH} protein initiates a repair event in which a cytosine base is inserted opposite the oxidized guanine base. In the absence of \textit{MUTYH} function, this form of DNA repair is defective, and an increase in guanine-to-thymine mutations is the result. The colonic epithelium is evidently more dependent on \textit{MUTYH} function than are most other cells in the body. Various reasons for the increased dependence have been cited, including the high proliferation rate of colonic stem cells, high levels of reactive oxygen species in the gut epithelium, and an overrepresentation of mutation hot spots in the \textit{APC} gene.\textsuperscript{13}

Mutations in the mismatch repair (MMR) genes \textit{MLH1}, \textit{MSH2}, \textit{MSH6}, and \textit{PMS2} are associated with Lynch syndrome, which is the most common hereditary CRC syndrome, accounting for as much as 3% of all CRCs.\textsuperscript{11} The MMR system removes mismatched bases that have been improperly incorporated into DNA during replication. Lynch syndrome is associated with a wide array of cancer sites, including the uterus, ovary, stomach, small intestine, bile ducts, ureter, and transitional cells in the kidney, brain, and skin.

Mutations in the \textit{APC}, \textit{MUTYH}, and MMR genes, together with several much less frequently mutated syndrome-causing genes, contribute to approximately 5% of all CRCs and constitute, at most, 40% of the total genetic variance in CRC.

\textbf{Figure 2}  \hspace{1em}  Cellular pathways dysregulated in colorectal cancer (CRC). Specific genetic factors that are altered in CRC and discussed in this review are blue. Figure altered from original by Wikipedia user RoadNotTaken (https://commons.wikimedia.org/wiki/File:Signal_transduction_pathways.svg, last accessed June 20, 2017). This image is being used with permission under the terms of the GNU Free Documentation License, version 1.2 or any later version, published by the Free Software Foundation (with no invariant sections, no front-cover texts, and no back-cover texts). The image herein originally appeared on November 18, 2010, and is current as of publication of this article. APC, adenomatous polyposis coli; BMP, bone morphogenetic protein; CDK, cyclin-dependent kinase; CREB, CAMP-responsive element-binding protein; EGF, epidermal growth factor; EPC, endothelial progenitor cell factors; ERK, extracellular signal–regulated kinase; FADD, Fas-associated protein with death domain; FasR, Fas receptor; GPCR, G protein–coupled receptor; GSK, glycogen synthase kinase; IGF, insulin-like growth factor; JAK, Janus-activating kinase; JNK, c-Jun N-terminal kinase; MAPK, mitogen-activated protein kinase; MEK, MAP kinase kinase; PI3K, phosphatidylinositol 3-kinase; PLC, phospholipase C; RSMAD, receptor phosphorylated SMAD; RTK, receptor tyrosine kinase; SARA, SMAD anchor for receptor activation; SMO, smoothed; SOS, son of sevenless; TCF, T-cell factor; TGF, transforming growth factor; Tnf, tumor necrosis factor.
risk. In the past 10 years, there have been attempts to close the gap between mendelian genetics and the genetic variance that remains to be explained via the conduct of genome-wide association studies (GWASs). GWASs compare the frequencies of common genetic variation—single-nucleotide polymorphisms (SNPs)—in CRC cases versus controls to identify SNPs associated with disease. These studies were successful in the identification of many new CRC risk factors across the genome; however, the proportion of the genetic variance in CRC that is explained by these newly identified risk factors is, at most, 5%. The effect sizes of these common genetic risk factors are relatively small (odds ratios, from 1.04 to 1.56); consequently, large studies (samples sizes in the hundreds of thousands) may be required to identify all of the common genetic risk factors. Recent work also focuses on genetic variation that is less common (allele frequencies, <5%) and on gene-environment interactions that might explain the so-called missing genetic variance.

It is against this backdrop of the rare, the common, and the unknown that we will consider the current state of affairs in the understanding of genetic risk factors for CRC in African Americans. As expected, African Americans are not exempt from occurrence of the rare hereditary syndromes. Early evidence showed that African Americans with hereditary forms of CRC harbored novel mutations in the MMR genes. Recent clinic-based reports have gone further to detail many more genetic mutations in APC and MUTYH and in the MMR genes in the African American population. Differences in the numbers or proportions of hereditary cases identified in African Americans compared with whites most likely reflect differences in the ordering of genetic testing rather than differences in the frequencies or the phenotypes associated with these mutations, although more extensive investigation of these questions is needed. Population-based studies have not addressed questions relating to the frequency of mutations in the known hereditary genes in African American populations, so that we do not know the impact of hereditary mutations on the incidence of CRC in African Americans and on the disparity in incidence. In the white population, hereditary mutations account for approximately 5% of all CRCs; thus, the difference in frequency of hereditary mutations would have to be large to explain the difference in incidence. Moreover, there is evidence against a large excess of MMR gene mutations—the largest class of hereditary mutations—based on phenotypic testing of African American CRCs (discussed in Similar Frequencies of Microsatellite Instability).

GWASs have identified >40 risk-associated regions on the basis of analysis of SNP data generated predominantly in European-ancestry CRC cases compared with controls. Theoretically, a higher frequency of common genetic risk factors predisposing to CRC in African Americans may explain differences in incidence between whites and African Americans. However, taking into account the differences in allele frequencies and odds ratios between the two populations, there is substantially more attributable risk estimated for the known risk-associated regions in European-ancestry populations compared with that in African Americans. This conclusion emanates from two studies, each including >1800 African American subjects, who were tested for replication of known risk-associated SNPs (Table 1). The studies found that slightly less than half of the known risk-associated regions were significantly associated with CRC in African Americans. African American populations have been included as part of larger GWASs (eg, in the identification of SNP associations in V711A). Because patterns of linkage disequilibrium differ between European- and African-ancestry populations, associations with SNPs in African Americans that are different from the index SNP in European-ancestry populations have been observed in several instances. In addition, novel associations within known risk-associated regions have been reported. Finally, using a GWAS design, a novel CRC African American-specific association with the SYMPK gene was identified.

Lack of replication for European-ancestry SNP associations has also been observed in an Asian CRC GWAS. Differences in allele frequencies and linkage disequilibrium structure between populations and population-specific gene-gene and gene-environment interactions have been proposed as explanations. Lack of adequate power to detect associations remains a problem in the African American GWASs. The observation of novel associations in which the risk-associated SNP is absent or present only at low frequencies in European-ancestry populations indicates the existence of population-specific or population-restricted risk alleles; many more of this kind of risk factor are likely to be encountered as investigations turn to alleles with frequencies of <5%. In summary, GWASs in African American CRC have uncovered a limited number of risk-associated regions, and additional GWASs are needed to fully explore the genetic architecture of CRC risk in African Americans. Analysis of risk alleles that have frequencies of <5% will face an additional obstacle from the higher frequencies of benign genetic variation in African-ancestry populations, making the distinction between signal and noise more challenging.

Can Vitamin D Levels Explain Differences in CRC Incidence in African Americans?

Vitamin D is a hormone; its production is initiated in basal keratinocytes in the skin as a result of exposure to UVB (280 to 320 nm) rays in sunlight. The cholecalciferol generated in the skin is converted to calcidiol in the liver, and calcidiol is converted to calcitriol, which is the active hormone, by the kidney and other tissues. Approximately 95% of the total vitamin D is produced endogenously as opposed to obtained through diet or dietary supplementation. Vitamin D has numerous physiological effects, including regulation of calcium homeostasis, immunity, insulin secretion, and blood pressure and carcinogenesis. In case-control and cohort studies, higher levels of serum vitamin D are associated with a
Carolina subjects. Two novel associations were reported for index SNP that showed a stronger association.

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|l|l|l|l|}
\hline
Gene region* & Chr & SNP ID & Position, bp & Allele & F(A) & OR & P value & Ref Notes \\
\hline
\textbf{LIN2C1655} & 1q41 & 12759486 & 222,066,536 & T & 0.58 & 0.86 & $1.5 \times 10^{-4}$ & 19 Fine-mapping SNP \\
\textbf{C5orf66} & 5q31.1 & rs647161 & 134499052 & A & 0.55 & 1.14 & 0.002 & 19 Index SNP \\
\textbf{SLC22A3} & 6q26-27 & rs7758229 & 160,840,252 & T & 0.11 & 1.08 & 0.056 & 19 Index SNP \\
\textbf{EIF3H} & 8q23.3 & rs16892766 & 117,630,683 & C & 0.13 & 1.17 & 0.0058 & 19 Index SNP \\
\textbf{CASC20} & 20p12.3 & rs9683267 & 128,413,305 & T & 0.13 & 0.87 & 0.029 & 19 Index SNP \\
\textbf{MYC} & 824.21 & rs9683267 & 128,413,305 & T & 0.12 & 0.87 & 0.21 & 20 Index SNP \\
\textbf{MYC} & 824.21 & rs7014346 & 128,424,792 & G & 0.39 & 1.05 & 0.27 & 19 Index SNP \\
\textbf{MYC} & 824.21 & rs7014346 & 128,424,792 & G & 0.39 & 1.15 & 0.06 & 20 Index SNP \\
\textbf{POLD3} & 11q13.4 & rs3824999 & 74,345,550 & G & 0.20 & 1.15 & 0.009 & 19 Index SNP \\
\textbf{GREM1} & 15q13.3 & rs16969681 & 32,993,111 & T & 0.13 & 1.16 & 0.01 & 19 Index SNP \\
\textbf{GREM1} & 15q13.3 & rs10318 & 33,025,979 & T & 0.03 & 1.45 & 0.04 & 20 Index SNP \\
\textbf{GREM1} & 15q13.3 & rs11632715 & 33,004,247 & A & 0.38 & 1.04 & 0.34 & 19 Index SNP \\
\textbf{GREM1} & 15q13.3 & rs11632715 & 33,004,247 & A & 0.29 & 2.26 & 0.002 & 20 Index SNP \\
\textbf{GREM1} & 15q13.3 & rs17816285 & 33,039,298 & G & 0.29 & 3.13 & $2 \times 10^{-4}$ & 21 Novel SNP association \\
\textbf{CDH1} & 16q22.1 & rs9929218 & 68,820,946 & A & 0.29 & 0.93 & 0.12 & 19 Index SNP \\
\textbf{CDH1} & 16q22.1 & rs1862748 & 68,832,943 & T & 0.20 & 0.82 & 0.023 & 20 Index SNP \\
\textbf{RHPN2} & 19q13.1 & rs7252505 & 33,575,064 & A & 0.62 & 0.91 & $1.8 \times 10^{-4}$ & 19 Fine-mapping SNP \\
\textbf{RHPN2} & 19q13.1 & rs113984415 & 33,555,034 & A & 0.19 & 0.13 & $8 \times 10^{-5}$ & 21 Novel SNP association \\
\textbf{CASC20} & 20p12.3 & rs961253 & 6,404,281 & A & 0.36 & 1.08 & 0.054 & 19 Index SNP \\
\textbf{CASC20} & 20p12.3 & rs961253 & 6,404,281 & A & 0.37 & 0.93 & 0.30 & 20 Index SNP \\
\hline
\end{tabular}
\caption{Associations Identified in Two Replication Studies that Tested for Association at Known Susceptibility Regions}
\end{table}

The studies by Wang et al.\textsuperscript{19} and Kupfer et al.\textsuperscript{20} included African American subjects comparing 1894 cases with 4703 controls and 795 cases with 985 controls, respectively. Wang et al.\textsuperscript{19} tested 21 risk-associated regions, and Kupfer et al.\textsuperscript{20} tested 10 risk-associated regions. The two studies share a small subset of North Carolina subjects. Two novel associations were reported for \textit{GREM1} and \textit{RHPN2} in the study by Kupfer et al.\textsuperscript{19} Fine-mapping SNP is an SNP correlated with the index SNP that showed a stronger association \textit{P} value. Unless otherwise stated, a log-additive model was used to estimate effect size and \textit{P} value.

\footnotesize{\textsuperscript{*Assignment of a gene to the association is based on proximity to the nearest risk-associated SNP.} \textsuperscript{1The association shown was observed for left-sided colorectal cancer only. No association was observed for all colorectal cancer.} \textsuperscript{2Recessive model.} \textsuperscript{3Allele, allele used as reference; Chr, chromosome; F(A), frequency of the reference allele; ID, identification; OR, odds ratio (not adjusted for local ancestry); P\textit{ value, P\textit{ value of association; Ref, article from which data were reproduced; SNP, single-nucleotide polymorphism.}}}

lower frequency of CRC,\textsuperscript{27,28} and administration of vitamin D can suppress CRC in animal models.\textsuperscript{29}

Because cutaneous melanin absorbs UVB wavelengths, individuals with darker skin colors require more sun exposure to achieve equivalent levels of serum vitamin D than individuals with lighter skin colors.\textsuperscript{30} One apparent consequence of this difference is that African Americans have substantially lower levels of serum vitamin D than their white counterparts, especially in northern parts of the United States. Consequently, it has been hypothesized that some part of the disparity in CRC incidence may be explained by lower vitamin D levels in African Americans.\textsuperscript{31,32} In addition, lower vitamin D levels may also be associated with the higher rate of CRC-specific mortality in African Americans.\textsuperscript{31,32}

Because of the link between colorectal adenoma and CRC, one expectation of this hypothesis is that vitamin D levels should be inversely associated with adenoma occurrence, but the data in favor of this expectation have been inconsistent.\textsuperscript{33–35} Inverse association with vitamin D levels may be restricted to advanced adenomas,\textsuperscript{36} which are more predictive of CRC risk. Another expectation is that supplementation with vitamin D should reduce the frequency of adenoma recurrence. Negative results from a large prevention trial, in which 1000 IU of vitamin D\textsubscript{3} was administered for 3 to 5 years, argue against a strong effect of vitamin D supplementation on adenoma recurrence.\textsuperscript{37} This prevention trial also challenged the commonly held view that calcium supplementation reduces adenoma recurrence, although the prevalence of obesity in the study population may have attenuated response. The study did not address whether longer periods or higher levels of supplementation would provide protection from CRC and the extent to which vitamin D reduces progression of CRC. Moreover, few African Americans, the population with the highest risk because of low levels of serum vitamin D, were included in it.

Evidence for a contribution to African American CRC of genetic risk factors in vitamin D pathway genes is mostly negative,\textsuperscript{38–40} consistent with the finding that genetic variance contributes little to the variance in total vitamin D levels.\textsuperscript{31} There is evidence that individuals with specific vitamin D receptor (\textit{VDR}) genotypes might respond favorably to vitamin D supplementation and, by extension, to higher levels of vitamin D, whatever the source.\textsuperscript{42} Moreover, there are differences between African- and European-ancestry individuals in the transcriptional response of colonic epithelial cells to active vitamin D hormone administered \textit{in vitro}, suggesting that ethnicity-dependent regulatory circuitry may play a role in response to vitamin D levels \textit{in vivo}.\textsuperscript{43} At a broader level, many of the genetic variants that are associated with differences in gene
Dietary Influences on CRC and the Gut Microbiome

Migration studies strongly support a dominant role for environmental factors in the etiology of CRC. Populations migrating from countries with a low prevalence of CRC to countries with a high prevalence of CRC acquire the higher CRC prevalence of their new country within a generation or less. Individual dietary components are often credited for increasing (e.g., red meat consumption) or decreasing (e.g., fiber) CRC risk, but the evidence is uneven and relatively weak. For example, individuals in the highest quartile of red meat consumption have an approximately 40% increased risk of CRC over individuals in the lowest quartile. This association was observed in retrospective studies but was not significant in prospective studies. African Americans have been underrepresented in most studies; however, observations made in the North Carolina Colon Cancer Study and a sister study administered by the same group that included rectal cancer cases suggest that dietary patterns may explain some of the racial differences in CRC.

Recently, attention has turned to the gut environment, specifically its microbial communities, to better understand the interplay between dietary factors and host microbiota in the development of CRC. The gut microbiota consists of trillions of microbes, the vast majority of which are bacteria. The term microbiome refers to the collective genome harbored by the microbiota. Microbiota-host interactions comprise a continuum of symbiosis, commensalism, and pathogenicity. Growing evidence implicates biofilms and adverse species of adherent bacteria in the process of CRC development. Broadly, there are three ways in which the microbiota may contribute to carcinogenesis: (i) by acting as cancer-initiating microbes (oncomicrobes), (ii) by guiding adverse immune system function, and (iii) by producing carcinogenic metabolites. There is solid evidence for all three mechanisms, but the relative contribution of each is unknown.

Incidence rates of CRC are vastly different for African Americans (60 per 100,000 per year) and South African blacks (5 per 100,000 per year). Of the many differences that characterize the environment for these different individuals, diet can play an outsized role in the incidence rates for CRC. The diet for rural South African blacks is highly enriched in fiber and low in meat and fat, whereas the Western diet is low in fiber and high in meat and fat. O’Keefe and colleagues conducted a diet switch in which they gave Pittsburgh African Americans the traditional South African black diet, whereas they gave rural South African blacks the Western diet. They then examined changes in the gut microbiome and fecal metabolites.

Reciprocal changes were observed in the gut microbiome and the metabolome within 2 weeks of diet switch. African Americans consuming the high-fiber, low-fat diet exhibited increases in the abundance of microbial species involved in fiber fermentation, including species that metabolize butyrate and other short-chain fatty acids and the hydrogenotrophic microbes that remove hydrogen, including methanogens, acetogens, and sulfate-reducing bacteria. Concomitantly, there was suppression of microbial species that process bile acids, including bile acid deconjugators and the pathogenic sulfidogenic bacteria and Fusobacterium nucleatum. South Africans receiving the low-fiber, high-fat diet exhibited the opposite pattern (i.e., a lower abundance of species engaged in saccharolytic fermentation and higher bile acid deconjugating species). In addition, there were also reciprocal changes in gut epithelial cell proliferation, which is a marker for cancer development. The secondary bile acids deoxycholic acid and lithocholic acid have long been hypothesized as carcinogenic, and they are metabolized from glycine- and taurine-conjugated primary bile acids by gut microbes. Consequently, dietary influences on the microbiome that affect production of secondary bile acids and other carcinogens are particularly important areas of future research.

One study has examined differences in the gut microbes in CRC cases between African Americans and whites. Yazici and colleagues measured the abundances of sulfidogenic bacteria, including sulfate-reducing bacteria and B. wadsworthia, in a well-characterized series of Chicago, IL, CRC cases and endoscopy clinic screening controls. Irrespective of disease status, African Americans exhibited much greater abundances of sulfidogenic bacteria than whites. In addition, African American CRC cases had higher levels of B. wadsworthia than African American controls, whereas controls had higher levels of sulfate-reducing bacteria. Although African Americans consumed higher levels of fat and protein than their white counterparts, only a relatively small proportion of the difference in microbial abundance were found to be accounted for by differences in dietary intake in the statistical models. The implication of these results is that an undisclosed racial factor(s) drives differences in microbial abundance.

Causal connections are being sought between obesity and all of the major human cancers, including CRC. Increased consumption of red meat and high-fat and high-glycemic index foods, coupled with reduced levels of physical activity, has fueled an epidemic of obesity and type 2 diabetes at ever-younger ages. African Americans have higher rates of diabetes than whites, more insulin resistance, and more hyperinsulinemia. On the other hand, measures of metabolic syndrome are paradoxically lower because African Americans have paradoxically lower levels of low-density lipoprotein cholesterol, higher high-density lipoprotein cholesterol, lower serum fatty acids, and lower triacylglycerol levels. Although it is not known whether these differences in pancreatic β cell and liver functions are...
environmentally or genetically driven (or both), they may play into the racial difference in CRC incidence. Further studies are needed to investigate the links among diabetes, obesity, and CRC in African Americans.

**Are Carcinogenic Mechanisms Different in African American CRC?**

In addition to risk of CRC, the frequencies of various clinicopathologic features of tumors that arise in African Americans are different from the frequencies found in whites. African Americans have a higher likelihood of a diagnosis in the proximal colon and have a higher prevalence of proximal adenomas than whites. These differences are important because, until recently, lesions in the proximal colon were more frequently missed during colonoscopy. African Americans also are often diagnosed at a younger age (median ages, 66 and 70 years for African American men and women compared with 72 and 77 years for white men and women, respectively). Moreover, African Americans are two times more likely to be diagnosed with CRC before the age of 50 years, which justified the recommendation to begin endoscopic screening at the age of 45 years instead of 50 years. The mechanisms driving early-onset CRC are not known, but a growing body of evidence suggests that they may be different from the mechanisms underlying the older age-of-onset CRCs. Although the field of carcinogenesis in African American CRC has advanced in its efforts to further investigate clinicopathologic differences, our understanding of the molecular pathogenesis of the CRCs that arise in this population remains in an early stage. However, recent articles have gone beyond the anecdotal evidence and raised questions about the progression of molecular events in CRCs in African Americans.

**Similar Frequencies of Microsatellite Instability**

A subset of CRCs is driven by a deficiency in DNA MMR genes, which causes a type of genomic instability known as microsatellite instability (MSI). Deficiency of an MMR gene results in increased genomic instability, particularly in microsatellite loci, leading to frequent changes in repeat sequence length in tumor DNA when compared with normal tissue. In the general population, MSI occurs in approximately 15% of CRCs, and it has hereditary and sporadic forms. Patients with Lynch syndrome carry a germline mutation in one of the MMR genes, and MSI develops when the remaining normal copy of the gene is mutated in a cancer precursor cell. The sporadic form of MSI usually develops as a result of DNA methylation of the *MLH1* promoter, and sporadic MSI is associated with more elderly female cases of CRC. Tumors that exhibit MSI have distinctive clinical features, including increased tumor-infiltrating lymphocytes and a Crohn disease—like inflammatory reaction; an association with right-sided disease; and an association with high-grade tumors with poorly differentiated, mucinous, or signet-ring phenotypes. Despite the higher grade of MSI tumors, affected individuals have a slightly better prognosis. Recently, it has been found that MSI tumors are more likely to respond to immune checkpoint therapy, likely because of the generation of excess neoantigens caused by the increased genomic instability.

Early reports suggested that African Americans had a higher frequency of CRC tumors with MSI. This was an interesting hypothesis because it had the potential to explain the higher frequency of right-sided CRC in African Americans. In furtherance of this idea, Eaton and colleagues studied a previously identified polymorphism in 5,10-methylenetetrahydrofolate reductase, whose effect on MSI status of a tumor was mediated by folate status, where the higher the folate intake, the lower the risk of an MSI tumor. Because this protective allele is less common in African Americans, if MSI frequency was increased in this population, then theoretically, it would contribute to an increased incidence of MSI. However, equal or lower frequencies of MSI in African American CRCs were observed in subsequent studies that had larger sample sizes.

A recent meta-analysis of 22 studies with a total of 12,611 CRC patients compared MSI frequencies in African Americans and whites. Conclusively, no difference in MSI frequencies was found. It is unknown the extent to which biological variables account for differences between studies. MSI is a relatively well-studied CRC subtype; however, it is unknown how much survival of patients with MSI differs by race, which may be important in addressing the health disparity in CRC mortality.

**Somatic Mutations in African American CRC**

Our understanding of molecular characteristics of microsatellite stable CRCs in African Americans has been lacking in the past because of the limited integration of these communities into population-level studies. This has improved in the past 5 years with studies focusing on chromosomal instability and single-nucleotide variants in African Americans.

Broad copy number variations in African Americans (ie, gains or losses that occur across chromosome arms or entire chromosomes) have been characterized in two studies, both of which have found frequencies of gains and losses to be similar to those seen in previous studies of primarily white patients. Frequent arm-level losses are seen in 1p, 8p, 14q, 15q, 18p, and 18q, whereas arm-level gains are frequently seen in 1q, 7p, 8q, 13q, 19q, 20p, and 20q. However, two other studies found that arm-level changes were different between the two populations. An unpublished study from our group (G.J.A. and N.A.E.) examined chromosome-arm copy number alterations but did not see qualitative differences; however, chromosome-arm gains were less frequent in African Americans. Adjustments for factors such as age and stage are needed to account for possible sampling bias in these studies. In addition, larger studies are necessary to determine the extent to
which clinicopathologic features can be explained by copy number variations.

Different focal alterations may exist in African Americans; however, larger platform-matched studies are necessary to determine a reliable list of affected genes. Despite these advances in exploring copy number variations in African American CRCs, data are still limited. Current studies have relied on small sample sizes (Varadan et al, American CRCs, data are still limited. Current studies have relied on small sample sizes (Varadan et al, n = 30 for sequence data and n = 12 for array data; Ashktorab and colleagues, n = 15), and they were not powered to account for differences in tumor location, stage, or differentiation status.

Genomic instability has also been noted in patient samples with elevated microsatellite alterations at selected tetranucleotide repeats (EMASTs). This EMAST instability phenotype has been described to be present in 60% of sporadic CRC cases, specifically in MSI, but also often in non-MSI cases. In a cohort of rectal cancer patients (n = 147, 26% African American), the EMAST phenotype was more frequent in African Americans, and it was associated with advanced stage. Because EMAST has been described in both MSI and microsatellite stable tumors, the etiological basis of EMAST may be heterogeneous. Further study is needed to understand the relationship of this instability to tumor progression and outcome in African American CRC.

DNA sequencing studies have been used to investigate somatic mutations in African American CRCs to determine the extent to which genes involved in tumor development are different in African American colorectal carcinogenesis. For the known cancer driver genes (KRAS, BRAF, TP53, PIK3CA, and APC) (Figure 2), frequencies of somatic mutations have been reported for relatively limited sample series (Table 2). Significant differences were not reported. Several groups have focused on the identification of novel driver genes that might explain features of the cancer health disparity. Ashktorab et al identified single-nucleotide variants in the CRC initiator APC that were unique to African Americans. Guda et al concentrated their efforts on identifying novel genes involved in advanced CRCs in African Americans. To do so, they sequenced 31 late-stage microsatellite stable CRCs from Case Medical Center (Cleveland, OH) and identified a 15-gene set that they followed up in a series of 129 African American cases in a replication stage. Most frequently mutated in this gene set was EPHA6, a member of the ephrin receptor tyrosine kinase family that has been previously implicated in CRC development. Another frequently mutated gene of interest was CHD5, a chromatin-remodeling protein that was also frequently mutated in a study by our group (G.J.A. and N.A.E., unpublished data). This gene set was used to determine differences in survival. It predicted poorer survival, especially in later-stage African American CRCs (n = 66). Large multi-institutional studies are needed to confirm and extend these results.

### Epigenetic Changes in African American CRC

There has been an increasing interest in changes to gene expression and DNA methylation in CRC, with several groups comparing African American CRCs with tumors from whites. An early study (n = 102, 50% African American) used methylation-specific PCR to look at methylation in promoters of a candidate cancer gene panel of 14 genes. Three genes were found to be hypermethylated in African Americans compared with white samples. Among them was CHD5, a tumor suppressor that was identified as a recurrently mutated gene that may be important in African American CRC development. This epigenetic silencing was further studied by Fatemi et al (n = 18 adenomas, and n = 6 healthy normal samples), who validated the down-regulation of CHD5 in adenoma samples and noted that this may be an early event in CRC tumor progression of African Americans. In addition, the promoters of ICAM5 (part of the intercellular adhesion molecule family) and GPNMB (a glycoprotein) were found to be hypermethylated in African American CRCs in the study by Mokarram et al.

Further studies have used more comprehensive technologies to discover more genes with DNA hypermethylation in African American CRC. The development of methylation microarrays, such as Illumina’s Infinium Human-Methylation27 and HumanMethylation450 BeadChip arrays (Illumina, San Diego, CA), and bisulfite sequencing techniques allow scientists to look genome wide for differential DNA methylation. Using microarray technology, Ashktorab and colleagues found 16 genes with consistent promoter
hypermethylation in a small cohort of African Americans (n = 12 CRCs, n = 8 adenomas, and n = 2 normal colon samples). Among these genes, ICAM5 and DCC were previously identified, which was studied by Mokarram et al\textsuperscript{100} but not found to be differentially methylated. In addition, the promoter of EVL was found to be more highly methylated in their African American CRC cohort than reported in other white cohorts. Well-known tumor suppressors were also included in this list, including APC and PTEN.

Reduced representation bisulfite sequencing gives an even more granular look at the epigenome by increasing the number of CpGs covered into the millions. Ashktorab et al\textsuperscript{103} sought to identify potential biomarkers for African American CRC by identifying novel hypermethylated genes using reduced representation bisulfite sequencing. This study elucidated four genes that were hypermethylated, including EID3, GPR75, GAS7, and BMP3. The latter two of these genes are in the insulin and transforming growth factor-β network, and other G-protein-coupled receptors are within this network as well. This interesting result connects tumorigenesis with insulin signaling that is affected in diabetes. Other studies have shown that diabetes can increase an individual’s risk for CRC\textsuperscript{94}, however, the biological basis of this association is not yet known. In another study from the same group,\textsuperscript{95} additional CpG islands were identified as hypermethylated in African American CRC, including L3MBTL1, NKK6-2, PREX1, TRAF7, PRDM14, and NEFM. This small study focused on increased methylation from normal to adenoma to malignancy (n = 5 CRCs, n = 2 adenomas, n = 1 blood, and n = 1 normal colon). These genes map to pathways often dysregulated during cancer development (namely, WNT/β-catenin, phosphatidylinositol 3-kinase, AKT, vascular endothelial growth factor, and Janus-activating kinase/STAT3 signaling pathways) (Figure 2). The methylation of these genes was specific to CRCs and not found in adenomas.

Another study\textsuperscript{96} took a comparative approach, comparing DNA methylation in African American and white CRCs (n = 6 African Americans, and n = 7 whites). Several differentially methylated genes were miRNAs, two of which (ie, miR-9 and miR-124) have previously been implicated in CRC development.\textsuperscript{97–99} Among the differentially methylated miRNAs was miR-124-3p, which was hypermethylated in African American CRCs compared with white CRCs. RNA-sequencing analysis found that two targets of this miRNA (ie, POLR2B and CYP1B1) were among the upregulated genes in these cases.

These studies are limited by their sample size of cancer patients, because most incorporated <10 samples into their genome-wide analyses. In addition, integration of methylation data with clinicopathologic and molecular features is needed. For example, it is known that DNA methylation and expression profiles differ in CRCs with MSI\textsuperscript{100}; therefore, stratifying analyses based on this criterion may remove some heterogeneity. Notably, there was no overlap in the genes identified in different studies, which raises the question of reproducibility.

**CRC-Specific Gene Dysregulation**

A recent landmark article based on European-ancestry CRC cases described four major expression-based CRC subtypes, referred to as consensus molecular subtypes (CMSs).\textsuperscript{100} CMS1 closely represents the MSI-immune CRCs, with BRAF mutations, hypermutation, a hypermethylation phenotype, and immune cell infiltration. CMS2 represents the canonical CRC subtype, driven by WNT signaling and MYC activation and containing excess copy number alterations. CMS3 represents metabolic CRC, driven by KRAS mutations and associated with metabolic dysregulation and higher genome stability. Finally, CMS4 represents mesenchymal CRC, associated with stromal infiltration, transforming growth factor-β activation, and angiogenesis markers. These subtypes are important because CMS1 is associated with worse survival after relapse and CMS4 is associated with worse relapse-free and overall survival.

In studies of African American CRC, this area experiences some of the same limitations (ie, reproducibility and covariate availability issues) as others when looking at its impact on the health disparity. Two studies compared gene expression in African American and white patients with CRCs.\textsuperscript{96,101} These studies identified 142 and 10 dysregulated genes, respectively, of which no genes were identified in both studies. The larger of the two studies, by Jovov and colleagues,\textsuperscript{101} validated expression of their 10 genes in an independent set of patients (n = 86, 50% African American). It would be interesting to reanalyze these data in light of the CMS subtype classification system,\textsuperscript{100} including the CMS1 subtype that is associated with MSI status.

In addition, two studies have focused on miRNA expression. Bovell et al\textsuperscript{102} (n = 106 African Americans, and n = 239 whites) found high miR-203 expression to be associated with poor survival in early-stage African American CRCs, whereas high miR-181b expression was associated with poor survival in stage III CRCs from African Americans. The prognostic value differed from the white patients in the study, suggesting that expression profiles may have a different prognostic value, depending on a patient’s race and the stage of the disease. Li et al\textsuperscript{103} found five miRNAs whose expression levels differed by race in their cohort (n = 30 African Americans, and n = 31 whites). Specifically, miR-182 was up-regulated in African American CRCs compared with white CRCs, and two potential miR-182 targets (FOXO1 and FOXO3A) were expressed at lower levels in African American CRCs.

Both expression of genes and expression of regulatory molecules, such as miRNAs, may act as possible prognostic biomarkers for CRCs in African Americans. Replication studies are necessary to validate the above findings. In addition, application of molecular pathological
epidemiology, integrating information from multiple sources, such as CMS status, clinicopathologic features, and environmental/behavioral factors, would result in better design of future studies.

The Continuing Problem of CRC Mortality in African Americans

As previously mentioned, African Americans have the highest mortality rate of CRC of any ethnic group in the United States. This disparity has been noted in previous reviews, which looked at behaviors, stage of tumor, treatment, and socioeconomic status as potential explanations. Some of this difference can be explained by tumor stage and socioeconomic status, but the disparity remains even after accounting for tumor stage at diagnosis and after adjusting for socioeconomics, comorbidities, and treatment. The effect of body mass index on survival was assessed, and although it had an effect on survival, it was determined not to contribute to the disparity.

The rate of mortality for CRC has decreased since the early 2000s; however, the mortality rate of African Americans is still 35% higher than that of whites (Figure 1B). The rate of decrease in mortality has improved as well, with both African Americans and whites having annual percentage changes in rate of -2.5% to 3%. Increased endoscopic screening by African Americans has contributed to the decrease in mortality.

Several biomarkers have been proposed to enable a more accurate prognosis for African American CRC patients. Some, including nuclear accumulation of p53 and mucin 1, have been unsuccessful. Others have been weak markers for prognosis, such as Bcl2 (whose lack of expression was weakly associated with survival, particularly in distal tumors) and p27kip1 (whose lowered expression was associated with worse survival in African Americans and whites but only in stage III CRCs).

Jones et al identified a polymorphism in the glutathione S-transferase gene (GSTP1) that is associated with a decreased risk of death for the genotype Ile/Val or Val/Val. This effect was more pronounced in patients who underwent chemotheraphy. However, the genotype frequency did not vary by race, and was determined not to contribute to the disparity.

Additional work was done to identify if somatic gene mutations may predict a CRC patient outcome. This study tested for KRAS, BRAF, and PIK3CA mutations as well as MSI status. They found that KRAS mutations were more frequent in African Americans and were associated with advanced stage in these patients. PIK3CA mutations were associated with decreased survival when adjusted for MSI, education, and income; however, the CI for the hazard ratio estimate was wide, limiting the interpretation of this piece of data. The authors suggest that the different frequency of somatic alterations may explain some of the differences in survival, a gap that needs to be addressed in future research.

Overall, despite research that has given us valuable knowledge about factors that have effects on CRC mortality, we still know relatively little about the molecular mechanisms underpinning why African Americans with CRC are more likely to die from the disease than other ethnic groups. More work is necessary to determine the effectiveness of current treatments in this population. In addition, differences in tumor progression markers may explain some of the differences in mortality; however, further research is needed to determine whether validated prognostic biomarkers can be developed.

References


Augustus and Ellis


Dawood SJ, Lippert MF, et al: The association between paternal high-fat diet and an increased risk of colorectal cancer in the FALCHS (Family, Lifestyle and Colorectal Health Study). PLoS One 2013, 8:e71176


The American Journal of Pathology