CLINICAL AT

Associations Between Colorectal Cancer Molecular Markers and Pathways With Clinicopathologic Features in Older Women

N. JEWEL SAMADDER,[1] ROBERT A. VIERKANT,[2] LORI S. TILLMANS,[3] ALICE H. WANG,[2] DANIEL J. WEISENBERGER,[4] PETER W. LAIRD,[4] CHARLES F. LYNCH,[5] KRISTIN E. ANDERSON,[6] AMY J. FRENCH,[3] ROBERT W. HAILE,[7] JOHN D. POTTER,[8] SUSAN L. SLAGER,[2] THOMAS C. SMYRK,[3] STEPHEN N. THIBODEAU,[3] JAMES R. CERHAN,[9] and PAUL J. LIMBURG[10]

[1]Department of Medicine (Gastroenterology), Huntsman Cancer Institute and University of Utah, Salt Lake City, Utah; [2]Division of Biomedical Statistics and Informatics and [3]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota; [4]USC Epigenome Center, Norris Comprehensive Cancer Center, Los Angeles, California; [5]Department of Epidemiology, University of Iowa, Iowa City, Iowa; [6]Department of Epidemiology, University of Minnesota, Minneapolis, Minnesota; [7]Department of Preventive Medicine, Keck School of Medicine of USC, Los Angeles, California; [8]Fred Hutchinson Cancer Research Center, Seattle, Washington; [9]Division of Epidemiology and [10]Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota

BACKGROUND & AIMS: Colorectal tumors have a large degree of molecular heterogeneity. Three integrated pathways of carcinogenesis (ie, traditional, alternate, and serrated) have been proposed, based on specific combinations of microsatellite instability (MSI), CpG island methylator phenotype (CIMP), and mutations in *BRAF* and *KRAS*. We used resources from the population-based Iowa Women's Health Study (n = 41,836) to associate markers of colorectal tumors, integrated pathways, and clinical and pathology characteristics, including survival times. **METHODS:** We assessed archived specimens from 732 incident colorectal tumors and characterized them as microsatellite stable (MSS), MSI high or MSI low, CIMP high or CIMP low, CIMP negative, and positive or negative for *BRAF* and/or *KRAS* mutations. Informative marker data were collected from 563 tumors (77%), which were assigned to the following integrated pathways: traditional (MSS, CIMP negative, *BRAF* mutation negative, and *KRAS* mutation negative; n = 170), alternate (MSS, CIMP low, *BRAF* mutation negative, and *KRAS* mutation positive; n = 58), serrated (any MSI, CIMP high, *BRAF* mutation positive, and *KRAS* mutation negative; n = 142), or unassigned (n = 193). Multivariable-adjusted Cox proportional hazards regression models were used to assess the associations of interest. **RESULTS:** Patients' mean age ($P = .03$) and tumors' anatomic subsite ($P = .0001$) and grade ($P = .0001$) were significantly associated with integrated pathway assignment. Colorectal cancer (CRC) mortality was not associated with the traditional, alternate, or serrated pathways, but was associated with a subset of pathway-unassigned tumors (MSS or MSI low, CIMP negative, *BRAF* mutation negative, and *KRAS* mutation positive) (n = 96 cases; relative risk = 1.76; 95% confidence interval, 1.07–2.89, compared with the traditional pathway). **CONCLUSIONS: We identified clinical and pathology features associated with molecularly defined CRC subtypes. However, additional studies are needed to determine how these features might influence prognosis.**

Keywords: Molecular Epidemiology; Colon Cancer; Prognostic Factor; Integrated Pathways.

Based on recent global estimates, colorectal cancer (CRC) is the third most common malignancy worldwide.[1] Although often viewed as a single disease, CRC more accurately represents a constellation of heterogeneous subtypes that result from different combinations of genetic events and epigenetic alterations. A growing body of evidence supports the ability to aggregate CRC subtypes based on combinations of microsatellite instability (MSI), CpG island methylator phenotype (CIMP), somatic *BRAF* mutation, and/or somatic *KRAS* mutation status.[2–11] For example, compared with MSS/MSI-low tumors, MSI-high tumors are more likely to be located in the proximal colon, lower stage, higher grade, and associated with increased tumor-infiltrating lymphocytes.[12,13] CIMP-positive (or CIMP-high) tumors have been associated with older age, proximal colonic location, poor differentiation and MSI-high status.[3,9,14,15] Outside of familial syndromes, somatic *BRAF* mutation (exon 15, V600E) appears to be strongly correlated with CIMP-positive or CIMP-high tumors.[3,16] Somatic *KRAS* mutations (particularly in codons 12 and 13) are reportedly more common in CIMP-positive tumors with a lesser degree of hypermethylation (CIMP low), and have also been associated with the MSS/MSI-low and *BRAF*-mutation–negative CRC subtypes.[3,17]

To further clarify the complex relationships among MSI, CIMP, *BRAF*, and *KRAS* status in colorectal carcinogenesis, several integrated molecular models have been described previously.[11,18–20] In a recent special issue of *Gastroenterology* dedicated to CRC updates and future directions, Leggett and Whitehall proposed the following predominant pathways for sporadic CRC development, building from existing integrated models and further incorporating the timing of critical molecular alterations[11]: the traditional pathway, characterized by early *APC* mutation and chromosomal instability, resulting in MSI-low or MSS, CIMP-negative, *BRAF*-mutation–negative, and *KRAS* mutation–negative tumors; the alternate pathway, in

Abbreviations used in this paper: CI, confidence interval; CIMP, CpG island methylator phenotype; CRC, colorectal cancer; IWHS, Iowa Women's Health Study; MSI, microsatellite instability; PCR, polymerase chain reaction; RR, relative risk; SEER, Surveillance, Epidemiology, and End Results.

0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2013.05.001

which either *KRAS* or *APC* mutation precedes development of MSI-low or MSS, CIMP-low tumors; and the serrated pathway, in which *BRAF* mutation can lead to CRCs with MSI-high, CIMP-high, or MSI-low or MSS, CIMP-high phenotype.

At present, anatomic extent of disease (as represented by TNM stage) is the most commonly employed measure for estimating CRC prognosis.[21,22] Yet, the TNM system[23] does not adequately account for within-stage, subhistologic heterogeneity, prompting recommendations for additional assessment of molecular markers as adjuncts to, or modifiers of, TNM staging.[21,22,24–26] To date, the traditional, alternate, and serrated pathways have not been characterized with respect to their clinicopathologic associations or prognostic potential in prospective, population-based studies. Data and tissue resources from the Iowa Women's Health Study (IWHS) of older women were used to generate novel data in this regard.

Materials and Methods

Approvals for the current study were obtained from the Institutional Review Boards for Human Research at Mayo Clinic Rochester, the University of Minnesota and the University of Iowa.

Cohort Recruitment and Case Ascertainment

Details regarding the methods used for recruitment and enrollment of IWHS participants have been reported elsewhere.[27] In brief, in January 1986, a 16-page baseline questionnaire was sent to 99,826 randomly selected women, ages 55–69 years, who resided in Iowa and held a valid driver's license. Of these, 41,836 women (42%) returned the baseline questionnaire, constituting the full IWHS subject cohort. Incident CRC cases were identified through the State Health Registry of Iowa, which participates in the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program.[28] Annual matching between a computer-generated list of all cohort members and the records of Iowans with incident cancer in the SEER program registry was performed based on combinations of first, last, and maiden names; ZIP code; birth date; and social security number. Demographic characteristics and CRC incidence rates for baseline survey responders and nonresponders have been shown to be similar, as reported previously.[29] Data on tumor location, grade, SEER stage, chemotherapy exposure, and radiation therapy exposure were obtained from the Iowa registry. CRCs located in the cecum, ascending colon, hepatic flexure, transverse colon, and splenic flexure (ICD-O codes 18.0, 18.2–18.5) were categorized as proximal colon and cancers located in the descending colon, sigmoid colon, rectosigmoid junction and rectum (ICD-O codes 18.6, 18.7, 19.9, 20.9) were categorized as distal colon or rectum.

Mortality Data

Vital status and state of residence for IWHS participants were determined by mailed questionnaires in 1987, 1989, 1992, 1997, and 2004, as well as through linkage to Iowa death certificate records. Nonrespondents to follow-up surveys were compared against the National Death Index to identify decedents. Previous studies have estimated that 99% of all cancer-related deaths among IWHS cohort members are captured through this approach.[30]

Tissue Collection and Processing

Archived, paraffin-embedded tissue specimens were requested for incident CRC cases diagnosed from January 1, 1986 through December 31, 2002. For each participant, the pathology laboratory of record was contacted through an introductory request letter, with a second request letter and additional telephone request as needed. Pathology reports, diagnostic slides, and tissue blocks were mailed directly to Iowa Cancer Registry staff for initial accessioning, followed by shipment to the study laboratory coordinator (LST) at Mayo Clinic Rochester. Confirmation of CRC diagnosis and tissue block selection were performed for each case by an experienced gastrointestinal pathologist (TCS). Tissue specimens were retrieved from 732 of 1255 (58%) incident CRC cases. Of note, similar or lower incident CRC case numbers and/or retrieval rates have been recently reported in molecular epidemiology studies embedded within other large cohorts, such as the Nurses' Health Study (n = 528 [58%])[31] and the Health Professionals Follow-Up Study (n = 438 [51%]).[32] To assess the possibility of selection bias, associations among subject demographics, exposure patterns, and tumor characteristics (size and stage) were compared between patients whose tissue specimens could be retrieved and those whose specimens could not be retrieved. No statistically significant differences were observed for any comparisons, as reported previously.[33] Tissue sections were serially cut in 5- or 10-μm thick increments. H&E staining was used to identify areas of tumor (ie, >50% dysplastic cells in field of view) and normal tissue. In total, pathology materials were retrieved for 732 incident CRC cases. Tissue samples were scraped from unstained slides and placed into separate tubes. DNA extraction was performed using the QIAamp tissue kit (Qiagen, Valencia, CA) in accordance with manufacturer's instructions. High-quality, usable DNA samples were available from 563 of 732 (77%) cases.

Molecular Markers

MSI status was determined from paired tumor and normal DNA samples using 10 established microsatellite markers: 4 mononucleotide repeats (BAT25, BAT26, BAT40, and BAT34C4), 5 dinucleotide repeats (ACTC, D5S346, D18S55, D17S250, and D10197), and 1 complex marker (MYCL). Polymerase chain reaction (PCR) for the various microsatellite markers was carried out under standard conditions (95°C for 12 min followed by 38 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s, with a final extension for 10 min at 72°C) with a master mix that included 10× buffer type II, *Taq* gold, and all 4 deoxyribonucleotide triphosphates. Primers were custom ordered with various fluorescent dyes from Applied Biosystems (Foster City, CA). PCR products were analyzed on an ABI 3100 (Applied Biosystems). MSI status was categorized as MSI high if at least 3 of 10 markers demonstrated instability, MSI low if 1 or 2 of 10 markers demonstrated instability, microsatellite stable (MSS) if 0 of 10 markers demonstrated instability, and MSI missing if assay results were noninformative/unavailable.

CIMP status was evaluated by treating tumor DNA with sodium bisulfite (Zymo Research, Orange, CA) and subsequently analyzed using an automated real-time, PCR-based MethyLight system, which quantitatively measures genome-specific DNA methylation levels in comparison with a methylated reference sample (M.*Sss*I-treated DNA) to calculate the percentage of