

Molecular Characteristics of Early-Onset Colorectal Cancer According to Detailed Anatomical Locations: Comparison With Later-Onset Cases

Tomotaka Ugai, MD, PhD^{1,2,*}, Koichiro Haruki, MD, PhD^{1,*}, Tabitha A. Harrison, MPH^{3,*}, Yin Cao, ScD^{4,5,*}, Conghui Qu, MS³, Andrew T. Chan, MD, MPH^{6,7,8,9}, Peter T. Campbell, PhD¹⁰, Naohiko Akimoto, MD, PhD¹, Sonja Berndt, PhD¹¹, Hermann Brenner, MD, MPH^{12,13,14}, Daniel D. Buchanan, PhD^{15,16,17}, Jenny Chang-Claude, PhD^{18,19}, Kenji Fujiyoshi, MD, PhD¹, Steven J. Gallinger, PhD²⁰, Marc J. Gunter, PhD²¹, Akihisa Hidaka, MD, PhD³, Michael Hoffmeister, MD, MSc¹², Li Hsu, PhD³, Mark A. Jenkins, PhD²², Roger L. Milne, PhD^{22,23,24}, Victor Moreno, PhD^{25,26,27,28}, Polly A. Newcomb, MPH, PhD^{3,29}, Reiko Nishihara, PhD^{1,30}, Rish K. Pai, PhD³¹, Lori C. Sakoda, MPH, PhD^{3,32}, Martha L. Slattery, MPH, PhD³³, Wei Sun, PhD³, Efrat L. Amitay, MPH, PhD¹², Elizabeth Alwers, MD, PhD¹², Stephen N. Thibodeau, PhD³⁴, Amanda E. Toland, PhD³⁵, Bethany Van Guelpen, PhD^{36,37}, Michael O. Woods, PhD³⁸, Syed H. Zaidi, PhD³⁹, John D. Potter, PhD³, Marios Giannakis, MD, PhD^{40,41}, Mingyang Song, MD, MSc^{2,4,5,30}, Jonathan A. Nowak, MD, PhD^{1,†}, Amanda I. Phipps, MPH, PhD^{3,29,†}, Ulrike Peters, MPH, PhD^{3,29,†} and Shuji Ogino, MD, PhD, MS^{1,2,41,42,†}

INTRODUCTION: Early-onset colorectal cancer diagnosed before the age of 50 years has been increasing. Likely reflecting the pathogenic role of the intestinal microbiome, which gradually changes across the entire colorectal length, the prevalence of certain tumor molecular characteristics gradually changes along colorectal subsites. Understanding how colorectal tumor molecular features differ by age and tumor location is important in personalized patient management.

¹Program in MPE Molecular Pathological Epidemiology, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA;

²Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA; ³Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, Washington, USA; ⁴Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St Louis, Missouri, USA; ⁵Division of Gastroenterology, Department of Medicine, Washington University School of Medicine, St. Louis, Missouri, USA; ⁶Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA; ⁷Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts, USA;

⁸Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA; ⁹Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA; ¹⁰Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York, USA; ¹¹Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; ¹²Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; ¹³Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany; ¹⁴German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany; ¹⁵Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, Victoria, Australia; ¹⁶University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria, Australia; ¹⁷Genetic Medicine and Family Cancer Clinic, The Royal Melbourne Hospital, Parkville, Victoria, Australia; ¹⁸Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany; ¹⁹University Medical Centre Hamburg-Eppendorf, University Cancer Centre Hamburg (UCCH), Hamburg, Germany; ²⁰Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada; ²¹Nutrition and Metabolism Section, International Agency for Research on Cancer, World Health Organization, Lyon, France; ²²Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia; ²³Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia; ²⁴Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia; ²⁵Oncology Data Analytics Program, Catalan Institute of Oncology (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain; ²⁶CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain; ²⁷Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain; ²⁸Colorectal Cancer Group, ONCOBELL Program, Bellvitge Biomedical Research Institute-IDIBELL, Hospitalet de Llobregat, Barcelona, Spain; ²⁹Department of Epidemiology, University of Washington, Seattle, Washington, USA; ³⁰Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA; ³¹Department of Laboratory Medicine and Pathology, Mayo Clinic Arizona, Scottsdale, Arizona, USA; ³²Division of Research, Kaiser Permanente Northern California, Oakland, California, USA; ³³Department of Internal Medicine, University of Utah, Salt Lake City, Utah, USA; ³⁴Division of Laboratory Genetics, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA; ³⁵Departments of Cancer Biology and Genetics and Internal Medicine, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio, USA; ³⁶Department of Radiation Sciences, Oncology Unit, Umeå University, Umeå, Sweden; ³⁷Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden; ³⁸Memorial University of Newfoundland, Discipline of Genetics, St. John's, Canada; ³⁹Ontario Institute for Cancer Research, Toronto, Ontario, Canada; ⁴⁰Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA; ⁴¹Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA; ⁴²Cancer Immunology and Cancer Epidemiology Programs, Dana-Farber Harvard Cancer Center, Boston, Massachusetts, USA.

Correspondence: Shuji Ogino, MD, PhD, MS. E-mail: sogino@bwh.harvard.edu.

Use of Standardized Official Symbols: We use HUGO (Human Genome Organization)-approved official symbols (or root symbols) for genes and gene products, including BRAF, EGFR, KRAS, MLH1, MSH2, MSH6, and PMS2, all of which are described at www.genenames.org. Gene symbols are italicized whereas symbols for gene products are not italicized.

*Tomotaka Ugai, Koichiro Haruki, Tabitha A. Harrison, Yin Cao contributed equally as cofirst authors.

†Jonathan A. Nowak, Amanda I. Phipps, Ulrike Peters, Shuji Ogino contributed equally as colast authors.

Received August 24, 2022; accepted December 16, 2022; published online December 30, 2022

METHODS: Using 14,004 cases with colorectal cancer including 3,089 early-onset cases, we examined microsatellite instability (MSI), CpG island methylator phenotype (CIMP), and *KRAS* and *BRAF* mutations in carcinomas of the cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum and compared early-onset cases with later-onset cases.

RESULTS: The proportions of MSI-high, CIMP-high, and *BRAF*-mutated early-onset tumors were lowest in the rectum (8.8%, 3.4%, and 3.5%, respectively) and highest in the ascending colon (46% MSI-high; 15% CIMP-high) or transverse colon (8.6% *BRAF*-mutated) (all $P_{\text{trend}} < 0.001$ across the rectum to ascending colon). Compared with later-onset tumors, early-onset tumors showed a higher prevalence of MSI-high status and a lower prevalence of CIMP-high status and *BRAF* mutations in most subsites. *KRAS* mutation prevalence was higher in the cecum compared with that in the other subsites in both early-onset and later-onset tumors ($P < 0.001$). Notably, later-onset MSI-high tumors showed a continuous decrease in *KRAS* mutation prevalence from the rectum (36%) to ascending colon (9%; $P_{\text{trend}} < 0.001$), followed by an increase in the cecum (14%), while early-onset MSI-high cancers showed no such trend.

DISCUSSION: Our findings support biogeographical and pathogenic heterogeneity of colorectal carcinomas in different colorectal subsites and age groups.

KEYWORDS: colorectal continuum; colorectal neoplasm; epigenetics; mismatch repair; molecular pathological epidemiology

SUPPLEMENTARY MATERIAL accompanies this paper at <http://links.lww.com/AJG/C852>

Am J Gastroenterol 2023;118:712–726. <https://doi.org/10.14309/ajg.0000000000002171>

INTRODUCTION

There is a growing concern on early-onset colorectal cancer diagnosed before the age of 50 years, incidence of which has increased worldwide since around 1990 (1). Evidence suggests that, compared with later-onset cases, early-onset colorectal cancers occur more frequently in rectal location and less frequently in proximal colon location (2). Studies also indicate possible heterogeneity of molecular characteristics between early-onset and later-onset colorectal cancers (1–3). For instance, compared with later-onset cases, early-onset colorectal cancers are more commonly microsatellite instability (MSI)–high and less commonly CpG island methylator phenotype (CIMP)–high and *BRAF*-mutated (3–10). Considering these findings, it is of particular interest to examine the molecular pathology of early-onset colorectal cancers in comparison with later-onset tumors according to tumor location.

Colorectal cancer consists of biologically heterogeneous neoplasms with differing sets of genetic and epigenetic alterations, influenced by the microbiome and immune system (11,12), which may at least partly explain variable tumoral characteristics according to tumor location (13). Despite the pathophysiological importance of luminal contents and the intestinal microbiota (which gradually change along the colorectal length) (14), numerous studies have used a dichotomy model of proximal (right-sided) vs distal (left-sided) colorectum and have shown the associations of proximal tumors with high-level MSI, high-level CIMP, and *BRAF* mutations (15–19). By contrast, fewer studies have examined tumor molecular features according to more detailed colorectal segments (with somewhat limited case numbers in each subsite) (13,20–22). Likely reflecting the pathogenic role of the intestinal microbiome, which gradually changes across the entire colorectal length (14), it is conceivable that the prevalence of certain tumor molecular features of early-onset colorectal cancer might gradually change along colorectal subsites. However, molecular

features of early-onset colorectal cancer according to detailed subsites remain to be studied.

This consortium pooled analysis was conducted to test our hypotheses that the prevalence of major molecular features in early-onset colorectal cancer might change along detailed colorectal locations and that the trend might differ from that of late-onset colorectal cancer. In addition, previous studies showed that cecal cancer had a higher prevalence of *KRAS*-mutated tumors than all other colorectal subsites (13,20). Hence, another hypothesis was that the association of cecal cancer with *KRAS* mutations might be different between early-onset and later-onset colorectal cancers. We used 14,004 cases with colorectal cancer including, 3,089 early-onset cases, derived from The Cancer Genome Atlas (TCGA) and participating studies in the Genetics and Epidemiology of Colorectal Cancer Consortium.

METHODS

Study population

We use the term “early onset” for colorectal cancer diagnosed before the age of 50 years and the contrasting term “later onset” for colorectal cancer diagnosed at or after 50 years of age. We pooled data for 14,004 cases of colorectal cancer with available data on tumor location and tumor molecular characteristics from the following 12 studies: the Colon Cancer Family Registry, the Cancer Prevention Study II (23), the German Darmkrebs: Chancen der Verhütung durch Screening Study (24), the Diet, Activity and Lifestyle Study (25), the Early Detection Research Network (26), the European Prospective Investigation into Cancer–Sweden (27), the Health Professionals Follow-up Study (28), the Melbourne Collaborative Cohort Study (29), the Newfoundland Familial Colorectal Cancer Registries (30), the Nurses’ Health Study (31), the Northern Sweden Health and Disease Study (32), and TCGA (33). These studies, except TCGA, have

Table 1. Patient characteristics of colorectal cancer according to primary tumor location

Clinical and molecular feature ^a	Total, N	Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectum	P value
All cases	14,004	2,142	1,962	1,390	1,039	3,310	4,161	
Sex								<0.001 ^b
Female	6,712 (48%)	1,129 (53%)	1,088 (55%)	704 (51%)	505 (49%)	1,559 (47%)	1,727 (42%)	
Male	7,292 (52%)	1,013 (47%)	874 (45%)	686 (49%)	534 (51%)	1,751 (53%)	2,434 (58%)	
Mean age ± SD	61.4 ± 13.1	63.4 ± 12.8	64.2 ± 13.3	62.7 ± 13.6	60.2 ± 13.3	61.2 ± 12.3	58.9 ± 13.0	<0.001 ^b
Age group								<0.001 ^b
Younger than 50 yr (early-onset)	3,089 (22%)	367 (17%)	327 (17%)	271 (19%)	264 (25%)	687 (21%)	1,173 (28%)	
Mean age ± SD	42.4 ± 6.0	42.5 ± 5.8	41.6 ± 6.7	41.3 ± 6.8	42.3 ± 6.3	42.9 ± 5.8	42.7 ± 5.6	<0.001 ^b
50 yr or older (later-onset)	10,915 (78%)	1,775 (83%)	1,635 (83%)	1,119 (81%)	775 (75%)	2,623 (79%)	2,988 (72%)	
Mean age ± SD	66.7 ± 8.9	67.7 ± 9.0	68.7 ± 9.0	67.9 ± 9.0	66.3 ± 8.9	66.0 ± 8.4	65.2 ± 9.0	<0.001 ^b
Year of diagnosis								<0.001 ^b
Before 1995	2,218 (16%)	413 (20%)	318 (17%)	260 (19%)	196 (19%)	673 (21%)	358 (8.9%)	
1995–2000	5,721 (42%)	943 (45%)	740 (39%)	588 (43%)	417 (41%)	1,325 (41%)	1,708 (42%)	
After 2000	5,725 (42%)	736 (35%)	860 (45%)	522 (38%)	401 (40%)	1,235 (38%)	1,971 (49%)	
Family history of colorectal cancer								<0.001 ^b
Absent	10,273 (76%)	1,553 (75%)	1,359 (72%)	968 (72%)	734 (73%)	2,504 (78%)	3,155 (79%)	
Present	3,222 (24%)	512 (25%)	522 (28%)	375 (28%)	269 (27%)	690 (22%)	854 (21%)	
AJCC disease stage								<0.001 ^b
I	3,279 (25%)	469 (23%)	433 (23%)	269 (21%)	192 (19%)	839 (28%)	1,077 (29%)	
II and III	8,228 (63%)	1,311 (65%)	1,247 (67%)	895 (69%)	671 (68%)	1,827 (60%)	2,277 (60%)	
IV	1,457 (11%)	248 (12%)	173 (9.3%)	136 (10%)	127 (13%)	361 (12%)	412 (11%)	
MSI status								<0.001 ^c
Non-MSI-high	10,903 (83%)	1,440 (72%)	1,156 (62%)	848 (67%)	829 (84%)	2,942 (94%)	3,688 (95%)	
MSI-high	2,193 (17%)	570 (28%)	695 (38%)	418 (33%)	154 (16%)	179 (5.7%)	177 (4.6%)	
CIMP status								<0.001 ^c
Low/negative	7,926 (84%)	1,103 (72%)	856 (62%)	680 (70%)	610 (90%)	2,120 (94%)	2,557 (95%)	
High	1,561 (16%)	422 (28%)	532 (38%)	286 (30%)	66 (9.8%)	128 (5.7%)	127 (4.7%)	
<i>BRAF</i>								<0.001 ^c
Wild-type	10,957 (89%)	1,549 (82%)	1,297 (74%)	946 (79%)	828 (92%)	2,791 (95%)	3,546 (97%)	
Mutant	1,396 (11%)	346 (18%)	454 (26%)	254 (21%)	73 (8.1%)	145 (4.9%)	124 (3.4%)	
<i>KRAS</i>								0.69 ^c
Wild-type	6,710 (67%)	874 (55%)	957 (67%)	699 (71%)	498 (67%)	1,727 (70%)	1,955 (68%)	
Mutant	3,369 (33%)	710 (45%)	479 (33%)	282 (29%)	248 (33%)	726 (30%)	924 (32%)	

Table 1. (continued)

Clinical and molecular feature ^a	Total, N	Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectum	P value
Tumor subtype ^d								
Type 1	512 (7.5%)	157 (15%)	222 (24%)	91 (15%)	17 (3.4%)	23 (1.3%)	2 (0.1%)	<0.001 ^c
Type 2	219 (3.2%)	50 (4.8%)	74 (8.0%)	40 (6.8%)	14 (2.8%)	29 (1.6%)	12 (0.6%)	<0.001 ^c
Type 3	2,193 (32%)	438 (42%)	286 (31%)	158 (27%)	169 (34%)	508 (29%)	634 (32%)	0.21 ^c
Type 4	3,663 (54%)	330 (31%)	268 (29%)	268 (45%)	281 (56%)	1,183 (67%)	1,333 (66%)	<0.001 ^c
Type 5	257 (3.8%)	77 (7.3%)	78 (8.4%)	34 (5.8%)	19 (3.8%)	21 (1.2%)	28 (1.4%)	<0.001 ^c

AJCC, American Joint Committee on Cancer; CIMP, CpG island methylator phenotype; MSI, microsatellite instability.
^aPercentage indicates the proportion of patients with a specific patient characteristic among all patients or in strata of tumor location (cecum, ascending colon, transverse colon, sigmoid colon, and rectum).
^bTo compare categorical data between subgroups classified by the tumor location, the χ^2 test was performed. To compare continuous variables, an analysis of variance was performed.
^cP value was calculated by the linear trend test across the ordinal categories of subsite location variable (from the rectum to ascending colon) in the multivariable logistic regression model adjusted for sex, family history of colorectal cancer (present vs absent), and study (i.e., cohort).
^dTumor subtypes described by Jass (38) as follows: type 1 = MSI-high, CIMP-high, BRAF, mutant; KRAS, wild-type; type 2 = non-MSI-high, CIMP-high, BRAF, wild-type; type 3 = non-MSI-high, CIMP-low/negative, BRAF, wild-type; type 4 = non-MSI-high, CIMP-low/negative, BRAF, wild-type; type 5 = MSI-high, CIMP-low/negative, BRAF, wild-type; KRAS, mutant; type 4 = non-MSI-high, CIMP-low/negative, BRAF, wild-type; type 5 = MSI-high, CIMP-low/negative, BRAF, wild-type; KRAS, wild-type.

participated in the Genetics and Epidemiology of Colorectal Cancer Consortium. All study participants provided informed consent, and each study was approved by their relevant research ethics committee or institutional review board. Details of the studies were described in the previous publication from this consortium (34). All cases with colorectal cancer included in each study were confirmed, and clinical and pathological data were extracted through review of medical records, pathological reports, and/or death certificates. Tumor location data were recorded using the *International Classification of Disease* across studies. To harmonize the tumor location data, we included the hepatic flexure into the transverse colon, the splenic flexure into the descending colon, and the rectosigmoid junction into the rectum. Hence, we examined 6 anatomical subsites (the cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum).

Table 1 lists pertinent clinical and pathological features in the combined data set. Descriptive characteristics of colorectal cancer cases in each study are shown in Supplementary Digital Content (see Supplementary Table 1, <http://links.lww.com/AJG/C852>). In this study, patients with at least 1 biological parent or sibling affected with colorectal cancer (at least by self-report) were considered to have positive (present) colorectal cancer family history. As a pooled analysis, in many of the included studies, we could not obtain further information on colorectal or other cancers in family members, including the age of cancer diagnosis, cancers of non-colorectal organs, and the number of affected family members.

Evaluation of tumor molecular characteristics

Molecular marker testing for MSI, CIMP, BRAF, and KRAS statuses was conducted by each study according to individual study protocols (18,35). Details of methods and references for tumor molecular testing are described in Supplementary Digital Content (see Supplementary Materials and Supplementary Tables 2 and 3, <http://links.lww.com/AJG/C852>). A small fraction of the cohorts used loss of mismatch repair proteins (MLH1, MSH2, MSH6, and/or PMS2) as an acceptable surrogate of MSI-high status; we used the standardized nomenclature of genes and gene products as recommended by the expert panel (36). We compared results of MSI, BRAF, and KRAS statuses from each study with results from centralized targeted sequencing. The tumor classifications from these 2 approaches were highly (more than 90%) concordant (37).

We also defined tumor subtypes according to the Jass classification scheme (18,38) as follows: type 1 (MSI-high, CIMP-high, BRAF mutant, KRAS wild-type); type 2 (non-MSI-high, CIMP-high, BRAF mutant, KRAS wild-type); type 3 (non-MSI-high, CIMP-low/negative, BRAF wild-type, KRAS-mutant); type 4 (non-MSI-high, CIMP-low/negative, BRAF wild-type, KRAS wild-type); and type 5 (MSI-high, CIMP-low/negative, BRAF wild-type, KRAS wild-type).

Statistical analyses

All statistical analyses were conducted using STATA software (version 15.1; Stata Corporation, College Station, TX), and all P values were 2-sided. We compared the prevalence (proportion among cases with colorectal carcinoma) of a given molecular subtype (MSI-high, CIMP-high, BRAF-mutated, KRAS-mutated, or each Jass subtype) in different colorectal subsites.

Our primary hypothesis testing was an assessment of a statistically significant trend in the prevalence of MSI-high, CIMP-

Table 2. Molecular characteristics of early-onset and later-onset colorectal cancers according to primary tumor location

Molecular feature ^a	Total N	Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectum	Multivariable OR (95% CI) ^b	<i>P</i> _{trend} ^b	<i>P</i> _{interaction} ^c
Age younger than 50 yr (early-onset)										
MSI status								1.88 (1.73–2.03)	<0.001	<0.001
Non-MSI-high	2,291 (79%)	222 (64%)	168 (54%)	137 (56%)	187 (75%)	569 (89%)	1,008 (91%)			
MSI-high	601 (21%)	123 (36%)	146 (46%)	106 (44%)	62 (25%)	67 (11%)	97 (8.8%)			
CIMP status								1.55 (1.28–1.88)	<0.001	0.002
Low/negative	1,058 (94%)	133 (86%)	93 (85%)	96 (93%)	96 (97%)	238 (96%)	402 (97%)			
High	71 (6.3%)	21 (14%)	16 (15%)	7 (6.8%)	3 (3.0%)	10 (4.0%)	14 (3.4%)			
<i>BRAF</i>								1.29 (1.14–1.46)	<0.001	<0.001
Wild-type	2,529 (95%)	310 (95%)	261 (92%)	212 (91%)	213 (95%)	580 (95%)	953 (96%)			
Mutant	133 (5.0%)	15 (4.6%)	23 (8.1%)	20 (8.6%)	12 (5.3%)	28 (4.6%)	35 (3.5%)			
<i>KRAS</i>								1.10 (1.01–1.20)	0.024	0.012
Wild-type	1,035 (66%)	107 (51%)	81 (59%)	83 (66%)	95 (68%)	260 (70%)	409 (70%)			
Mutant	532 (34%)	102 (49%)	57 (41%)	43 (34%)	44 (32%)	111 (30%)	175 (30%)			
Age 50 yr or older (later-onset)										
MSI status								2.15 (2.05–2.26)	<0.001	
Non-MSI-high	8,612 (84%)	1,218 (73%)	988 (64%)	711 (70%)	642 (87%)	2,373 (95%)	2,680 (97%)			
MSI-high	1,592 (16%)	447 (27%)	549 (36%)	312 (30%)	92 (13%)	112 (4.5%)	80 (2.9%)			
CIMP status								1.97 (1.88–2.07)	<0.001	
Low/negative	6,868 (82%)	970 (71%)	763 (60%)	584 (68%)	514 (89%)	1,882 (94%)	2,155 (95%)			
High	1,490 (18%)	401 (29%)	516 (40%)	279 (32%)	63 (11%)	118 (5.9%)	113 (5.0%)			
<i>BRAF</i>								1.92 (1.82–2.02)	<0.001	
Wild-type	8,428 (87%)	1,239 (79%)	1,036 (71%)	734 (76%)	615 (91%)	2,211 (95%)	2,593 (97%)			
Mutant	1,263 (13%)	331 (21%)	431 (29%)	234 (24%)	61 (9.0%)	117 (5.0%)	89 (3.3%)			
<i>KRAS</i>								0.99 (0.96–1.03)	0.67	
Wild-type	5,675 (67%)	767 (56%)	876 (67%)	616 (72%)	403 (66%)	1,467 (70%)	1,546 (67%)			
Mutant	2,837 (33%)	608 (44%)	422 (33%)	239 (28%)	204 (34%)	615 (30%)	749 (33%)			

CI, confidence interval; CIMP, CpG island methylator phenotype; MSI, microsatellite instability; OR, odds ratio.

^aPercentage indicates the proportion of patients with a specific patient molecular characteristic among all patients or in strata of tumor location (cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum).

^bMultivariable odds ratio (OR) (with 95% 95% CI) for molecular marker positivity (i.e., MSI-high, CIMP-high, *BRAF* mutation, or *KRAS* mutation), and *P*_{trend} were calculated by the linear trend test across the ordinal categories of subsite location variable (from rectum [1] to ascending colon [5]) in the multivariable logistic regression model adjusted for sex (female vs male), family history of colorectal cancer (present vs absent), and study (i.e., cohort).

^c*P*_{interaction} was calculated using the Wald test for the cross product of the ordinal subsite variable and age (younger than 50 years vs 50 years or older) in the multivariable logistic regression model that adjusted for sex (female vs male), family history of colorectal cancer (present vs absent), and study (i.e., cohort).

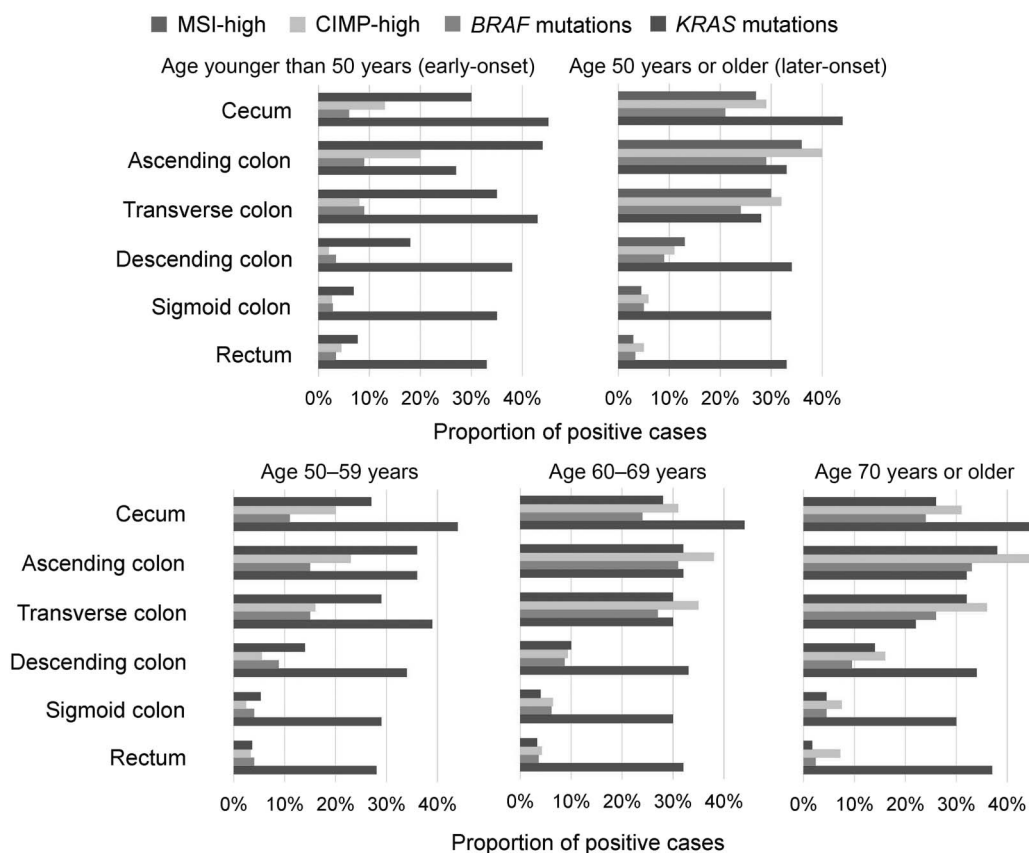


Figure 1. Prevalence of MSI-high status, CIMP-high status, *BRAF* mutations, and *KRAS* mutations along subslocations by age groups. CIMP, CpG island methylator phenotype; MSI, microsatellite instability.

high, or *BRAF*-mutated tumors along the detailed colorectal subslocations in early-onset and later-onset colorectal cancers. We calculated the multivariable-adjusted odds ratio (OR) of molecular marker positivity (with its corresponding P_{trend}) for 1 subsite-unit increase from the rectum to ascending colon (the ordinal categories of the rectum [1] to ascending colon [5]) in a logistic regression model with a given molecular marker as an outcome variable, adjusted for sex (female vs male), family history of colorectal cancer (present vs absent), and study (i.e., cohort). Missing values for family history of colorectal cancer ($N = 509$) were treated as separate indicator variables in the logistic regression model. We also tested another primary hypothesis that the trend of molecular marker prevalence from the rectum to ascending colon differed between early-onset and later-onset colorectal cancers. We used the Wald test for an interaction term between the subsite variable (ordinal; the rectum to ascending colon) and age (younger than 50 years vs 50 years or older). Because there were 12 primary hypotheses (4 marker trends in each of early-onset and later-onset groups and an interaction test between each marker [of 4 markers]) and age groups, we used the stringent 2-sided α level of 0.005 ($=0.05/12$) considering multiple comparisons (39).

In secondary analyses, we assessed the difference in *KRAS* mutation prevalence between cecum and other subsites, using multivariable logistic regression models (with the cecum location variable; yes vs no) that adjusted for sex (female vs male), family history of colorectal cancer (present vs absent), and study (i.e., cohort). Other secondary analyses included comparisons of

the proportions of molecular alterations between early-onset and later-onset cases (by the χ^2 tests) in selected colorectal subsites. We also assessed the relationships of the tumor location with each of the other variables listed in Table 1, using the χ^2 tests (or analysis of variance assuming equal variances for continuous variables).

RESULTS

Table 1 summarizes the characteristics of 14,004 cases with colorectal cancer including 3,089 early-onset (diagnosed before the age of 50 years) and 10,915 later-onset patients (diagnosed at or after age 50 years) according to primary tumor location. The proportion of rectal cancer was higher in early-onset cases (38%) than in later-onset cases (27%).

We examined statuses of MSI, CIMP, and *KRAS* and *BRAF* mutations in early-onset and later-onset colorectal cancers according to detailed subslocations (Table 2, Figure 1). The proportions of MSI-high, CIMP-high, and *BRAF*-mutated early-onset tumors were lowest in the rectum (8.8%, 3.4%, and 3.5%, respectively) and highest in the ascending colon (46% MSI-high; 15% CIMP-high) or transverse colon (8.6% *BRAF*-mutated) (all $P_{\text{trend}} < 0.001$ across the rectum to ascending colon), followed by declines in the cecum for MSI-high (36%) and *BRAF* mutation (4.6%) but not much for CIMP-high (14%). Similar increasing trends in the prevalence of MSI-high, CIMP-high, and *BRAF*-mutated tumors from the rectum to ascending colon were observed in later-onset colorectal cancers. Further age-stratified results on later-onset tumors are shown in Figure 1 and

Table 3. Molecular characteristics of early-onset and later-onset colorectal cancers according to primary tumor location in strata of sex

Molecular feature ^a	Total N	Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectum	Multivariable OR (95% CI) ^b	P _{trend} ^b
Age younger than 50 yr (early-onset)									
Female									
MSI status								1.90 (1.69–2.15)	<0.001
Non-MSI–high	1,230 (83%)	114 (70%)	93 (56%)	81 (65%)	110 (82%)	351 (93%)	481 (92%)		
MSI-high	257 (17%)	50 (30%)	73 (44%)	44 (35%)	24 (18%)	26 (6.9%)	40 (7.7%)		
CIMP status								1.68 (1.29–2.19)	<0.001
Low/negative	549 (93%)	65 (87%)	43 (80%)	49 (92%)	50 (98%)	151 (97%)	191 (96%)		
High	39 (6.6%)	10 (13%)	11 (20%)	4 (7.5%)	1 (2.0%)	4 (2.6%)	9 (4.5%)		
<i>BRAF</i>								1.43 (1.19–1.71)	<0.001
Wild-type	1,316 (95%)	153 (94%)	138 (91%)	112 (91%)	114 (97%)	344 (97%)	455 (97%)		
Mutant	64 (4.6%)	9 (5.6%)	14 (9.2%)	11 (8.9%)	4 (3.4%)	10 (2.8%)	16 (3.4%)		
<i>KRAS</i>								1.00 (0.89–1.13)	0.94
Wild-type	524 (63%)	50 (47%)	54 (73%)	36 (57%)	47 (62%)	146 (65%)	191 (67%)		
Mutant	305 (37%)	57 (53%)	20 (27%)	27 (43%)	29 (38%)	80 (35%)	92 (33%)		
Male									
MSI status								1.85 (1.66–2.05)	<0.001
Non-MSI–high	1,061 (76%)	108 (60%)	75 (51%)	56 (47%)	77 (67%)	218 (84%)	527 (90%)		
MSI-high	344 (24%)	73 (40%)	73 (49%)	62 (53%)	38 (33%)	41 (16%)	57 (10%)		
CIMP status								1.43 (1.06–1.94)	0.020
Low/negative	509 (94%)	68 (86%)	50 (91%)	47 (94%)	46 (96%)	87 (94%)	211 (98%)		
High	32 (5.9%)	11 (14%)	5 (9.1%)	3 (6.0%)	2 (4.2%)	6 (6.5%)	5 (2.3%)		
<i>BRAF</i>								1.20 (1.01–1.42)	0.037
Wild-type	1,213 (95%)	157 (96%)	123 (93%)	100 (92%)	99 (93%)	236 (93%)	498 (96%)		
Mutant	69 (5.4%)	6 (3.7%)	9 (6.8%)	9 (8.3%)	8 (7.1%)	18 (7.3%)	19 (3.7%)		
<i>KRAS</i>								1.24 (1.10–1.41)	0.001
Wild-type	511 (69%)	57 (56%)	27 (42%)	47 (75%)	48 (76%)	114 (79%)	218 (72%)		
Mutant	227 (31%)	45 (44%)	37 (58%)	16 (25%)	15 (24%)	31 (21%)	83 (28%)		
Age 50 yr or older (later-onset)									
Female									
MSI status								2.17 (2.03–2.33)	<0.001
Non-MSI–high	3,791 (79%)	593 (67%)	512 (60%)	314 (62%)	286 (84%)	1,050 (95%)	1,036 (97%)		
MSI-high	987 (21%)	298 (33%)	345 (40%)	193 (38%)	56 (16%)	58 (5.2%)	37 (3.4%)		

Table 3. (continued)

Molecular feature ^a	Total N	Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectum	Multivariable OR (95% CI) ^b	<i>P</i> _{trend} ^b
CIMP status								2.18 (2.03–2.34)	<0.001
Low/negative	3,065 (77%)	485 (66%)	380 (52%)	257 (59%)	246 (87%)	839 (94%)	858 (95%)		
High	921 (23%)	247 (34%)	357 (48%)	176 (41%)	36 (13%)	56 (6.3%)	49 (5.4%)		
<i>BRAF</i>								2.03 (1.89–2.18)	<0.001
Wild-type	3,713 (82%)	601 (72%)	515 (63%)	341 (69%)	272 (88%)	967 (94%)	1,017 (96%)		
Mutant	831 (18%)	237 (28%)	305 (37%)	150 (31%)	38 (12%)	57 (5.6%)	44 (4.1%)		
<i>KRAS</i>								0.90 (0.85–0.94)	<0.001
Wild-type	2,703 (67%)	436 (60%)	538 (73%)	312 (75%)	193 (68%)	637 (69%)	587 (63%)		
Mutant	1,315 (33%)	293 (40%)	201 (27%)	106 (25%)	91 (32%)	286 (31%)	338 (37%)		
Male									
MSI status								2.14 (1.99–2.31)	<0.001
Non-MSI-high	4,821 (89%)	625 (81%)	476 (70%)	397 (77%)	356 (91%)	1,323 (96%)	1,644 (97%)		
MSI-high	605 (11%)	149 (19%)	204 (30%)	119 (23%)	36 (9.2%)	54 (3.9%)	43 (2.5%)		
CIMP status								1.76 (1.64–1.90)	<0.001
Low/negative	3,803 (87%)	485 (76%)	383 (71%)	327 (76%)	268 (91%)	1,043 (94%)	1,297 (95%)		
High	569 (13%)	154 (24%)	159 (29%)	103 (24%)	27 (9.2%)	62 (5.6%)	64 (4.7%)		
<i>BRAF</i>								1.78 (1.65–1.93)	<0.001
Wild-type	4,715 (92%)	638 (87%)	521 (81%)	393 (82%)	343 (94%)	1,244 (95%)	1,576 (97%)		
Mutant	432 (8%)	94 (13%)	126 (19%)	84 (18%)	23 (6.3%)	60 (4.6%)	45 (2.8%)		
<i>KRAS</i>								1.09 (1.04–1.15)	<0.001
Wild-type	2,972 (66%)	331 (51%)	338 (60%)	304 (70%)	210 (65%)	830 (72%)	959 (70%)		
Mutant	1,522 (34%)	315 (49%)	221 (40%)	133 (30%)	113 (35%)	329 (28%)	411 (30%)		

CI, confidence interval; CIMP, CpG island methylator phenotype; MSI, microsatellite instability; OR, odds ratio.

^aPercentage indicates the proportion of patients with a specific patient molecular characteristic among all patients or in strata of tumor location (cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum).

^bMultivariable odds ratio (OR) (with 95% CI) for molecular marker positivity (i.e., MSI-high, CIMP-high, *BRAF* mutation, or *KRAS* mutation), and *P*_{trend} were calculated by the linear trend test across the ordinal categories of subsite location variable (from rectum (1) to ascending colon (5)) in the multivariable logistic regression model adjusted for family history of colorectal cancer (present vs absent) and study (i.e., cohort).

Table 4. Molecular characteristics of early-onset and later-onset colorectal cancers according to primary tumor location in strata of MSI status

Molecular feature ^a	Total N	Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectum	Multivariable OR (95% CI) ^b	<i>P</i> _{trend} ^b
Age younger than 50 yr (early-onset)									
Non-MSI-high									
CIMP status								1.55 (1.23–1.96)	<0.001
Low/negative	863 (95%)	93 (89%)	56 (84%)	56 (97%)	74 (97%)	215 (96%)	369 (97%)		
High	48 (5.3%)	11 (11%)	11 (16%)	2 (3.4%)	2 (2.6%)	9 (4.0%)	13 (3.4%)		
<i>BRAF</i>								1.47 (1.27–1.69)	<0.001
Wild-type	1,875 (95%)	189 (95%)	135 (88%)	105 (89%)	146 (94%)	478 (95%)	822 (97%)		
Mutant	104 (5.3%)	11 (6%)	18 (12%)	13 (11%)	9 (5.8%)	27 (5.3%)	26 (3.1%)		
<i>KRAS</i>								1.11 (0.99–1.24)	0.060
Wild-type	813 (67%)	68 (45%)	42 (53%)	48 (70%)	78 (72%)	224 (71%)	353 (70%)		
Mutant	408 (33%)	82 (55%)	37 (47%)	21 (30%)	30 (28%)	90 (29%)	148 (30%)		
MSI-high									
CIMP status								1.19 (0.72–1.94)	0.50
Low/negative	167 (89%)	39 (81%)	35 (88%)	36 (90%)	20 (95%)	11 (92%)	26 (96%)		
High	21 (11%)	9 (19%)	5 (13%)	4 (10%)	1 (4.8%)	1 (8.3%)	1 (3.7%)		
<i>BRAF</i>								1.05 (0.72–1.54)	0.80
Wild-type	478 (97%)	102 (98%)	115 (97%)	83 (95%)	52 (95%)	53 (100%)	73 (96%)		
Mutant	15 (3.0%)	2 (1.9%)	3 (2.5%)	4 (5.5%)	3 (5.5%)	0 (0%)	3 (3.9%)		
<i>KRAS</i>								1.04 (0.83–1.30)	0.74
Wild-type	139 (64%)	33 (69%)	34 (65%)	25 (61%)	13 (57%)	13 (68%)	21 (62%)		
Mutant	78 (36%)	15 (31%)	18 (35%)	16 (39%)	10 (43%)	6 (32%)	13 (38%)		
Age 50 yr or older (later-onset)									
Non-MSI-high									
CIMP status								1.63 (1.53–1.74)	<0.001
Low/negative	6,121 (91%)	802 (84%)	614 (78%)	468 (84%)	460 (93%)	1,765 (96%)	2,012 (95%)		
High	632 (9.4%)	157 (16%)	173 (22%)	90 (16%)	35 (7.1%)	80 (4.3%)	97 (4.6%)		
<i>BRAF</i>								1.61 (1.50–1.73)	<0.001
Wild-type	7,253 (94%)	1,011 (92%)	763 (86%)	551 (87%)	537 (94%)	2,038 (96%)	2,353 (97%)		
Mutant	498 (6.4%)	91 (8%)	129 (14%)	84 (13%)	33 (5.8%)	83 (3.9%)	78 (3.2%)		
<i>KRAS</i>								1.12 (1.08–1.16)	<0.001
Wild-type	4,378 (63%)	455 (46%)	438 (55%)	381 (66%)	336 (64%)	1,351 (70%)	1,417 (67%)		
Mutant	2,553 (37%)	542 (54%)	365 (45%)	199 (34%)	186 (36%)	567 (30%)	694 (33%)		
MSI-high									
CIMP status								1.63 (1.41–1.88)	<0.001
Low/negative	399 (34%)	108 (33%)	113 (26%)	70 (30%)	30 (52%)	43 (56%)	35 (88%)		
High	768 (66%)	217 (67%)	319 (74%)	165 (70%)	28 (48%)	34 (44%)	5 (13%)		
<i>BRAF</i>								1.57 (1.38–1.78)	<0.001
Wild-type	709 (50%)	179 (44%)	212 (42%)	138 (50%)	52 (67%)	69 (70%)	59 (92%)		
Mutant	716 (50%)	230 (56%)	288 (58%)	137 (50%)	26 (33%)	30 (30%)	5 (7.8%)		

Table 4. (continued)

Molecular feature ^a	Total N	Cecum	Ascending colon	Transverse colon	Descending colon	Sigmoid colon	Rectum	Multivariable OR (95% CI) ^b	<i>P</i> _{trend} ^b
<i>KRAS</i>								0.67 (0.57–0.79)	<0.001
Wild-type	1,017 (86%)	288 (86%)	401 (91%)	198 (85%)	52 (84%)	53 (71%)	25 (64%)		
Mutant	167 (14%)	47 (14%)	39 (9%)	35 (15%)	10 (16%)	22 (29%)	14 (36%)		

CI, confidence interval; CIMP, CpG island methylator phenotype; MSI, microsatellite instability; OR, odds ratio.

^aPercentage indicates the proportion of patients with a specific patient molecular characteristic among all patients or in strata of tumor location (cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum).

^bMultivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., CIMP-high, BRAF mutation, or KRAS mutation), and *P*_{trend} was calculated by the linear trend test across the ordinal categories of subsite location variable (from rectum (1) to ascending colon (5)) in the multivariable logistic regression model adjusted for sex (female vs male), family history of colorectal cancer (present vs absent), and study (i.e., cohort).

Supplementary Digital Content (see Supplementary Table 4, <http://links.lww.com/AJG/C852>).

In addition, we tested a hypothesis that the trend of molecular markers from the rectum to ascending colon differed between early-onset and later-onset tumors. The trends of the prevalence of MSI-high, CIMP-high, and BRAF-mutated tumors according to detailed sublocations significantly differed between early-onset and later-onset tumors (all *P*_{interaction} ≤ 0.001). Notably, compared with later-onset tumors, early-onset tumors showed a higher prevalence of MSI-high and a lower prevalence of CIMP-high and BRAF mutations in nearly all subsites (except for BRAF-mutated rectal tumors).

The proportion of KRAS-mutated early-onset tumors was higher in the cecum (49%) than in the other subsites (30%–41%) (multivariable OR for the cecum vs other subsites combined, 2.12 [95% confidence interval, 1.57–2.86]; *P* < 0.001). In later-onset tumors, cecal cancers showed a higher prevalence of KRAS mutation (44%) than cancers of other subsites (28%–33%) (the corresponding multivariable OR, 1.75 [95% confidence interval, 1.56–1.97]; *P* < 0.001). Stratified analyses by sexes and family history of colorectal cancer are summarized in Table 3, Supplementary Digital Content (see Supplementary Table 5, and Supplementary Figures 1–2, <http://links.lww.com/AJG/C852>). We also examined Jass tumor subtype classifications (38) according to detailed sublocations (see Supplementary Table 6, <http://links.lww.com/AJG/C852>). The proportions of type 1, 2, and 5 tumors increased from the rectum to ascending colon.

We further conducted analyses stratified by tumor characteristics (Table 4, Figure 2, and see Supplementary Tables 7–9, <http://links.lww.com/AJG/C852>). In early-onset non-MSI-high cases, the proportion of BRAF-mutated tumors increased from the rectum to ascending colon (*P*_{trend} < 0.001). In later-onset cases of both MSI-high and non-MSI-high, we observed continuous increases in the proportions of CIMP-high and BRAF-mutated tumors from the rectum to ascending colon (all *P*_{trend} < 0.001). Remarkably, later-onset MSI-high tumors showed a continuous decrease in KRAS mutation prevalence from the rectum (36%) to ascending colon (9%; *P*_{trend} < 0.001), followed by an increase in the cecum (14%). By contrast, early-onset MSI-high tumors did not show such a trend (*P*_{interaction} = 0.038, for the rectum-ascending colon trend in early-onset vs later-onset MSI-high cases). In addition, compared with later-onset MSI-high tumors, early-onset MSI-high tumors showed a lower prevalence of CIMP-high in all subsites. We further conducted stratified MSI-high cases by family history of colorectal cancer (see Supplementary Table 10, <http://links.lww.com/AJG/C852>). Although the sample size was limited, our findings tended to be consistent regardless of family history of colorectal cancer. Last,

we conducted stratified analysis by year of diagnosis (up to 2000 vs after 2000) and observed a similar trend in both strata (see Supplementary Table 11, <http://links.lww.com/AJG/C852>).

DISCUSSION

Colorectal adenocarcinoma represents a heterogeneous group of complex multifactorial diseases with varying cellular molecular features influenced by local tissue microenvironment. In this consortium pooled analysis using 3,089 early-onset and 10,915 later-onset cases, we found that the proportions of MSI-high, CIMP-high, and BRAF-mutated early-onset tumors were generally highest in the transverse and ascending colon and lowest in the rectum. In addition, compared with later-onset tumors, early-onset tumors showed a higher prevalence of MSI-high and a lower prevalence of CIMP-high and BRAF mutations in most subsites. The prevalence of KRAS mutation in both early-onset and later-onset tumors was consistently highest in the cecum. Notably, later-onset MSI-high tumors showed a continuous decrease in KRAS mutation prevalence from the rectum to ascending colon, followed by an increase in the cecum, but such a trend was not observed in early-onset MSI-high tumors. To our best knowledge, this study is the largest to date to investigate tumor molecular characteristics of early-onset and later-onset colorectal cancers along the detailed colorectal subsites.

Incidence of early-onset colorectal cancers in many organs diagnosed before the age of 50 years has globally been rising in recent decades for unknown reasons (40). Notably, most of early-onset cancer types that have shown the recent rise relate to the digestive system, potentially implicating the etiological role of the intestinal microbiota (40). Differences in the stool microbiome have been reported between patients with early-onset and later-onset colorectal cancer (41). The intestinal microbiota, which likely gradually changes across the entire colorectal length as luminal contents move toward the rectum, has been hypothetically linked with the continuous changes in the prevalence of tumor molecular features along the colorectal segments (13,14). Early-onset colorectal cancer has been associated with certain risk factor profiles, rectal location, signet ring cell histology, and specific tumor characteristics such as high-level MSI, LINE-1 hypomethylation, low/negative CIMP, BRAF wild-type, and lower tumor-infiltrating lymphocytes (1,3–7,42–46). This study attests to biogeographical (colorectal subsite) heterogeneity in molecular pathological features between early-onset and later-onset cases and even among early-onset cases. This study also provides rationale for the multisegmental design in the research of early-onset and later-onset colorectal cancer.

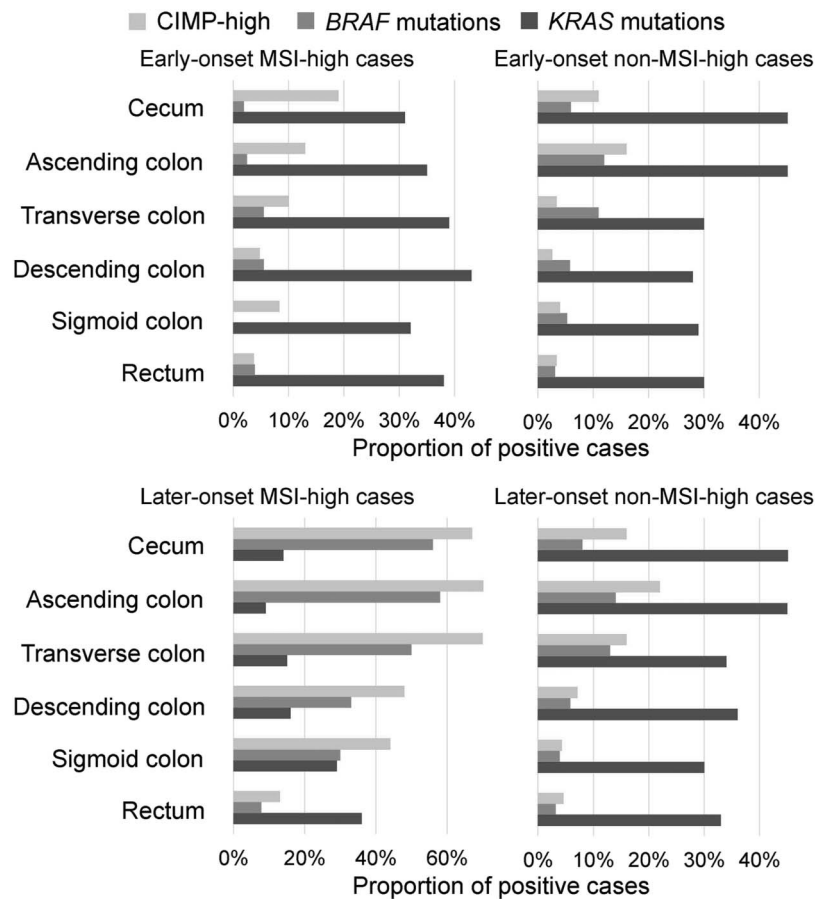


Figure 2. Prevalence of CIMP-high status, *BRAF* mutations, and *KRAS* mutations along sublocations by MSI status in early-onset and later-onset cancers. CIMP, CpG island methylator phenotype; MSI, microsatellite instability.

The proximal-distal dichotomy design has prevailed in colorectal disease research for decades and provided evidence for differences in molecular pathology between proximal and distal colorectal tumors (17). However, this dichotomy approach cannot provide insight into tumor characteristics in relation to more detailed subsites. On the contrary, previous studies have demonstrated differences in etiologies, molecular pathologies, and prognostic associations between the detailed anatomical subsites (13,20–22,47–49). Our current findings further support the importance of the colorectal multisegmental approach in clinical, epidemiological, and pathological research (13,50). This study also indicates that a large sample size is needed to examine tumors in each colorectal subsite with adequate statistical power.

In the colorectal tumor microenvironment, there is a dynamic interactive network that encompasses microorganisms and neoplastic, immune, and other cells, all of which are likely influenced by diet, lifestyle, environmental exposures, and other host factors (12,51). Accumulating evidence indicates that the gut microbiota may influence the pathogenesis of colorectal cancer through cellular molecular alterations and modulation of tumor immune interactions (52,53). An observational study has linked a certain dietary pattern with the abundance of sulfur-metabolizing bacteria in stool and shown that the pattern is associated with an increased risk of distal and rectal cancers (54). Another study suggests that the so-called Western dietary pattern rich in red and processed meat is associated with an increased risk of colorectal cancer, particularly in the distal

colon and rectum (55). Colonic epithelial cells are constantly in contact with bowel contents, including food debris, microorganisms, and their metabolites. Bowel contents and the gut microbiota likely vary across detailed anatomical subsites in the colorectum (14,56). Furthermore, we also showed that tumor immune microenvironment differed by early-onset and later-onset colorectal cancers (42). Hence, available pieces of evidence underscore the importance of the multisegmental approach in research on colorectal diseases including early-onset and later-onset colorectal cancer.

We found that later-onset MSI-high tumors exhibited a continuous decrease in *KRAS* mutation prevalence from the rectum to ascending colon, followed by an increase in the cecum, whereas early-onset MSI-high tumors did not show such a trend. MSI-high tumors are associated with distinct clinical and pathological features such as proximal tumor location, vigorous immune response, and better prognosis (57,58).

In addition, there exists heterogeneity in clinical and pathological features between early-onset and later-onset cases (1,3,42–46). Early-onset MSI-high tumors tend to be related to Lynch syndrome with germline mismatch repair gene mutations, while later-onset MSI-high tumors tend to be related to CIMP-high tumors with *MLH1* promoter hypermethylation (1,59). The intriguing difference in the *KRAS* mutation prevalence trend across colorectal subsites between early-onset and later-onset MSI-high cancers further emphasize pathobiological heterogeneity in colorectal cancer, which necessitates further investigations.

Cecal cancers seem to represent a unique subgroup of colorectal cancer, characterized by a high prevalence of *KRAS* mutations compared with cancers of the other sites in both early-onset and later-onset tumors, in agreement with the previous studies (13,20,60). Although the exact mechanism remains uncertain, the uniqueness of cecal cancers compared with cancers in the other colorectal segments may possibly reflect the following facts (1): that the cecum is the first place where luminal contents enter into the large intestine (2); that it has a pouch-like shape with a tendency of its content retention; and (3) that it is in close proximity to the ileum and appendix, both of which are rich in lymphoid immune tissue. Of note, appendiceal cancers have been shown to exhibit a high prevalence of *KRAS* mutation similar to proximal colon cancers (61,62) (also similar to cecal cancers in the previous study (13)). Thus, both cecal and appendiceal cancers are characterized by a high prevalence of *KRAS* mutations. Further studies are needed to elucidate biogeographical characteristics of the cecum distinct from the other colorectal segments.

We acknowledge the limitations of this study. First, we relied on tumor location information derived from individual medical records. Nonetheless, recording of colorectal tumor location has been standardized through the use of the *International Classification of Disease* code for each individual study. Second, tumor molecular analyses were performed using variable protocols across studies in this pooled analysis. However, all the molecular pathological methods have been well established with good validity in the literature of colorectal cancer research. Moreover, in a subset of cases, centralized targeted sequencing has shown high concordance of molecular pathological statuses for MSI, *KRAS*, and *BRAF*. Third, our patients were predominantly non-Hispanic Whites. Therefore, future analyses need to be conducted in different populations. Fourth, although we adjusted for multiple comparisons, false-positive findings could not be excluded. In addition, statistically significant but small differences may not be clinically important. Fifth, data on Lynch syndrome were not available. However, we stratified all cases and MSI-high cases by family history of colorectal cancer (see Supplementary Tables 6 and 10, <http://links.lww.com/AJG/C852>). Our findings tended to be consistent regardless of family history of colorectal cancer. Last, this analysis included older cases diagnosed before 2000, which may not be the same as a contemporary cohort, given the changing incidence trend of early-onset colorectal cancer. However, stratified analyses by year of diagnosis (up to 2000 vs after 2000) yielded similar trends in both strata (see Supplementary Table 11, <http://links.lww.com/AJG/C852>).

This study has unprecedented strengths. First and foremost, our consortium pooled analysis design with the large sample size enabled us to robustly evaluate the prevalence of the major tumor molecular features within each of the 6 subsites in strata of age, sex, colorectal cancer family history, and tumor molecular biomarker status. Second, cases with colorectal cancer in this study were drawn from several countries and based on different study designs, including case-control studies, prospective cohort studies, and multi-institutional case series, which likely increased generalizability.

In conclusion, this study showed substantial differences of tumor molecular characteristics in early-onset and later-onset colorectal carcinomas arising in different colorectal subsites. Our findings not only support biogeographical heterogeneity along colorectal length that influences carcinogenic processes but also provide compelling rationale for designing large-scale studies to robustly investigate detailed subsite data in colorectal disease research.

ACKNOWLEDGEMENTS

The Colon Cancer Family Registry (CCFR): The Colon CFR graciously thanks the generous contributions of their study participants, dedication of study staff, and the financial support from the US National Cancer Institute, without which this important registry would not exist.

CPS-II: We thank the CPS-II participants and Study Management Group for their invaluable contributions to this research. We acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries and cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

DACHS: We thank all participants and cooperating clinicians and Ute Handte-Daub, Utz Benscheld, Muhabbet Celik, and Ursula Eilber for excellent technical assistance.

EDRN: We acknowledge all the following contributors to the development of the resource: University of Pittsburgh School of Medicine, Department of Gastroenterology, Hepatology and Nutrition: Lynda Dzubinski; University of Pittsburgh School of Medicine, Department of Pathology: Michelle Bisceglia; and University of Pittsburgh School of Medicine, Department of Biomedical Informatics.

EPIC: Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization.

Harvard cohorts (HPFS, NHS): The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health and those of participating registries as required. We thank the participants and staff of the HPFS and NHS for their valuable contributions and the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, and WY. The authors assume full responsibility for analyses and interpretation of these data.

NSHDS: NSHDS investigators thank the Västerbotten Intervention Programme, the Northern Sweden MONICA study, the Biobank Research Unit at Umeå University, and Biobanken Norr at Region Västerbotten for providing data and samples and acknowledge the contribution from Biobank Sweden, supported by the Swedish Research Council.

SCCFR: We thank the study participants and staff of the Seattle Cancer Family Registry and the Hormones and Colon Cancer study (CORE Studies).

CONFLICTS OF INTEREST

Guarantor of the article: Shuji Ogino, MD, PhD, MS.

Specific author contributions: A.I.P., U.P., and S.O. and T.A.H.: conceived of the original study concept and design; K.H., T.U., and S.O.: designed the analyses; statistical analyses were performed by K.H., T.U., and C.Q.; K.H., T.U., and S.O. was assisted in the interpretation of results by Y.C., A.T.C., P.T.C., K.F., M.S., J.A.N., M.G., D.D.B., A.I.P., U.P., T.A.H.; K.H., T.U., and S.O. drafted the manuscript and all authors provided critical revisions to the manuscript for important intellectual content; T.A.H., A.T.C., P.T.C., S.B., H.B., D.D.B., J.C.-C., S.J.G., M.J.G., M.H., M.A.J., R.L.M., V.M., P.A.N., R.K.P., L.C.S., M.L.S., S.N.T., A.E.T., B.V.G., M.O.W., S.H.Z., J.D.P., M.S., A.I.P., U.P., and S.O. contributed to the acquisition of study data; P.A.N., M.L.S., M.O.W., A.I.P., U.P., and S.O. obtained funding contributing to this manuscript; T.A.H. and U.P. provided administrative, technical, or material support. Study supervision was provided by A.I.P., U.P., S.O., and T.A.H.

Financial support: Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO): National Cancer Institute, National Institutes of Health, US Department of Health and Human Services (U01 CA137088 and R01 CA248857). This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA015704. Scientific Computing Infrastructure at Fred Hutch is funded by ORIP grant S10OD028685. The Colon Cancer Family Registry (CCFR): The Colon Cancer Family Registry (CCFR, www.coloncfr.org) is supported in part by funding from the US National Cancer Institute (NCI), National Institutes of Health (NIH) (award U01 CA167551). The Seattle Colon Cancer Family Registry was also supported in part by the National Cancer Institute (NCI) of the US National Institutes of Health (NIH) under R01 CA076366 (to P.A.N.). The content of this manuscript does not necessarily reflect the views or policies of the US NCI, NIH, or any of the collaborating centers in the Colon Cancer Family Registry (CCFR) nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government, any cancer registry, or the CCFR. CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study-II (CPS-II) cohort. This study was conducted with Institutional Review Board approval. DACHS: This work was supported by the German Research Council (BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, HO 5117/2-1, HE 5998/2-1, KL 2354/3-1, RO 2270/8-1, and BR 1704/17-1), the Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A and 01ER1505B). DALs: the US National Institutes of Health (R01 CA48998 to M.L.S.). EDRN: This work is funded and supported by the US NCI, EDRN Grant (U01 CA 84968-06). EPIC: The coordination of EPIC is financially supported by the European Commission (DG SANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Éducation Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum, and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRC/Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), and Statistics Netherlands (The Netherlands); ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence programme on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); and Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom). Harvard Cohorts (HPFS, NHS): HPFS is supported by the US National Institutes of Health (P01 CA055075, U01 CA167552, U01 CA167552, R01 CA137178, R01 CA151993, R00 CA215314, and R35 CA197735) and NHS by the US National Institutes of Health (R01 CA137178, P01 CA087969, U01 CA186107, R01 CA151993, and

R35 CA197735). M.G. and S.O. were supported in part by the Cancer Research UK Grand Challenge Award (UK C10674/A27140). T.U. was supported by Prevent Cancer Foundation. MCCS: Melbourne Collaborative Cohort Study (MCCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 396414, and 1074383 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database. New South Wales (NSW) cancer registry data were obtained through the ACD with the assistance of the NSW Ministry of Health. NFCCR: This work was supported by an Interdisciplinary Health Research Team award from the Canadian Institutes of Health Research (CRT 43821); the National Institutes of Health, US Department of Health and Human Services (U01 CA74783); and National Cancer Institute of Canada grants (18223 and 18226). We acknowledge the contribution of Alexandre Belisle and the genotyping team of the McGill University and Génome Québec Innovation Centre, Montréal, Canada, for genotyping the Sequenom panel in the NFCCR samples. Funding was provided to M.O.W. by the Canadian Cancer Society Research Institute. NSHDS: This work was supported by Biobank Sweden through funding from the Swedish Research Council (VR 2017-00650, VR 2017-01737), the Swedish Cancer Society (CAN 2017/581), Region Västerbotten (VLL-841671, VLL-833291), Knut and Alice Wallenberg Foundation (VLL-765961), and the Lion's Cancer Research Foundation (several grants) and Insamlingsstiftelsen, both at Umeå University.

Potential competing interests: A.T.C. previously served as a consultant for Bayer Healthcare and Pfizer Inc. M.G. receives research funding from Bristol-Myers Squibb and Merck. A.T.C. is a Stuart and Suzanne Steele MGH research scholar. The other authors declare no potential conflicts of interest. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US NIH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Study Highlights

WHAT IS KNOWN

- ✓ The incidence of early-onset colorectal cancer diagnosed before the age of 50 years has increased worldwide.
- ✓ Early-onset colorectal cancer commonly occurs in the rectum.
- ✓ Early-onset colorectal cancer has tumor characteristics different from later-onset colorectal cancer.
- ✓ Tumor characteristics of colorectal cancer differ by primary tumor location.

WHAT IS NEW HERE

- ✓ Molecular characteristics of early-onset colorectal cancer changed gradually along colorectal subsites.
- ✓ Compared with later-onset tumors, early-onset tumors showed higher microsatellite instability–high prevalence and lower CpG island methylator phenotype–high/*BRAF* mutation prevalence in most subsites.
- ✓ Tumor molecular features varied by both age at diagnosis and detailed tumor location.

REFERENCES

- Akimoto N, Ugai T, Zhong R, et al. Rising incidence of early-onset colorectal cancer—a call to action. *Nat Rev Clin Oncol* 2021;18(4):230–43.
- Archambault AN, Su YR, Jeon J, et al. Cumulative burden of colorectal cancer-associated genetic variants is more strongly associated with early-onset vs late-onset cancer. *Gastroenterology* 2020;158(5):1274–86.e12.
- Willauer AN, Liu Y, Pereira AAL, et al. Clinical and molecular characterization of early-onset colorectal cancer. *Cancer* 2019;125(12):2002–10.
- Akimoto N, Zhao M, Ugai T, et al. Tumor long interspersed nucleotide element-1 (LINE-1) hypomethylation in relation to age of colorectal cancer diagnosis and prognosis. *Cancers (Basel)* 2021;13(9):2016.
- Antelo M, Balaguer F, Shia J, et al. A high degree of LINE-1 hypomethylation is a unique feature of early-onset colorectal cancer. *PLoS One* 2012;7(9):e45357.
- Baba Y, Huttenhower C, Noshho K, et al. Epigenomic diversity of colorectal cancer indicated by LINE-1 methylation in a database of 869 tumors. *Mol Cancer* 2010;9(1):125.
- Lieu CH, Golemis EA, Serebriiskii IG, et al. Comprehensive genomic landscapes in early and later onset colorectal cancer. *Clin Cancer Res* 2019;25(19):5852–8.
- Stoffel EM, Koeppe E, Everett J, et al. Germline genetic features of young individuals with colorectal cancer. *Gastroenterology* 2018;154(4):897–905.e1.
- Pearlman R, Frankel WL, Swanson B, et al. Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. *JAMA Oncol* 2017;3(4):464–71.
- Liang JT, Huang KC, Cheng AL, et al. Clinicopathological and molecular biological features of colorectal cancer in patients less than 40 years of age. *Br J Surg* 2003;90(2):205–14.
- Yamauchi M, Lochhead P, Morikawa T, et al. Colorectal cancer: A tale of two sides or a continuum? *Gut* 2012;61(6):794–7.
- Inamura K, Hamada T, Bullman S, et al. Cancer as microenvironmental, systemic and environmental diseases: Opportunity for transdisciplinary microbiomics science. *Gut* 2022;71(10):2107–22.
- Yamauchi M, Morikawa T, Kuchiba A, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut* 2012;61(6):847–54.
- Mima K, Cao Y, Chan AT, et al. *Fusobacterium nucleatum* in colorectal carcinoma tissue according to tumor location. *Clin Transl Gastroenterol* 2016;7(11):e200.
- Ogino S, Noshho K, Kirkner GJ, et al. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut* 2009;58(1):90–6.
- Chen Y, Chen Z, Huang J, et al. Clinicopathological and molecular characteristics of early-onset vs late-onset colorectal cancer according to tumor location. *Int J Clin Oncol* 2022;27(4):749–55.
- Lee MS, Menter DG, Kopetz S. Right versus left colon cancer biology: Integrating the consensus molecular subtypes. *J Natl Compr Canc Netw* 2017;15(3):411–9.
- Phipps AI, Alwers E, Harrison T, et al. Association between molecular subtypes of colorectal tumors and patient survival, based on pooled analysis of 7 international studies. *Gastroenterology* 2020;158(8):2158–68.e4.
- Domingo E, Freeman-Mills L, Rayner E, et al. Somatic POLE proofreading domain mutation, immune response, and prognosis in colorectal cancer: A retrospective, pooled biomarker study. *Lancet Gastroenterol Hepatol* 2016;1(3):207–16.
- Loree JM, Pereira AAL, Lam M, et al. Classifying colorectal cancer by tumor location rather than sidedness highlights a continuum in mutation profiles and consensus molecular subtypes. *Clin Cancer Res* 2018;24(5):1062–72.
- Rosty C, Young JP, Walsh MD, et al. PIK3CA activating mutation in colorectal carcinoma: Associations with molecular features and survival. *PLoS One* 2013;8(6):e65479.
- Phipps AI, Buchanan DD, Makar KW, et al. BRAF mutation status and survival after colorectal cancer diagnosis according to patient and tumor characteristics. *Cancer Epidemiol Biomarkers Prev* 2012;21(10):1792–8.
- Campbell PT, Deka A, Briggs P, et al. Establishment of the cancer prevention study II nutrition cohort colorectal tissue repository. *Cancer Epidemiol Biomarkers Prev* 2014;23(12):2694–702.
- Brenner H, Chang-Claude J, Seiler CM, et al. Protection from colorectal cancer after colonoscopy: A population-based, case-control study. *Ann Intern Med* 2011;154(1):22–30.
- Slattery ML, Friedman GD, Potter JD, et al. A description of age, sex, and site distributions of colon carcinoma in three geographic areas. *Cancer* 1996;78(8):1666–70.
- Crichton DJ, Mattmann CA, Thornquist M, et al. Bioinformatics: Biomarkers of early detection. *Cancer Biomark* 2011;9(1-6):511–30.
- Riboli E, Kaaks R. The EPIC project: Rationale and study design. European prospective investigation into cancer and nutrition. *Int J Epidemiol* 1997;26(90001):S6–14.
- Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med* 2013;369(12):1095–105.
- Giles GG, English DR. The Melbourne collaborative cohort study. *IARC Sci Publ* 2002;156:69–70.
- Stuckless S, Parfrey PS, Woods MO, et al. The phenotypic expression of three MSH2 mutations in large Newfoundland families with Lynch syndrome. *Fam Cancer* 2007;6:1–12.
- Liao X, Lochhead P, Nishihara R, et al. Aspirin use, tumor PIK3CA mutation, and colorectal-cancer survival. *N Engl J Med* 2012;367(17):1596–606.
- Van Guelpen B, Hultdin J, Johansson I, et al. Low folate levels may protect against colorectal cancer. *Gut* 2006;55(10):1461–6.
- The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;487(7407):330–7.
- Peters U, Jiao S, Schumacher FR, et al. Identification of genetic susceptibility loci for colorectal tumors in a genome-wide meta-analysis. *Gastroenterology* 2013;144(4):799–807.e24.
- Hidaka M, Eguchi S, Okuda K, et al. Impact of anatomical resection for hepatocellular carcinoma with microportal invasion (vp1): A multi-institutional study by the kyushu study group of liver surgery. *Ann Surg* 2020;271(2):339–46.
- Fujiyoshi K, Bruford EA, Mroz P, et al. Opinion: Standardizing gene product nomenclature—a call to action. *Proc Natl Acad Sci USA* 2021;118(3):e2025207118.
- Zaidi SH, Harrison TA, Phipps AI, et al. Landscape of somatic single nucleotide variants and indels in colorectal cancer and impact on survival. *Nat Commun* 2020;11(1):3644.
- Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* 2007;50(1):113–30.
- Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav* 2018;2(1):6–10.
- Ugai T, Sasamoto N, Lee HY, et al. Is early-onset cancer an emerging global epidemic? Current evidence and future implications. *Nat Rev Clin Oncol* 2022;19(10):656–73.
- Kong C, Liang L, Liu G, et al. Integrated metagenomic and metabolomic analysis reveals distinct gut-microbiome-derived phenotypes in early-onset colorectal cancer. *Gut Published Online First: 11 August 2022*.
- Ugai T, Vayrynen JP, Lau MC, et al. Immune cell profiles in the tumor microenvironment of early-onset, intermediate-onset, and later-onset colorectal cancer. *Cancer Immunol Immunother* 2022;71(4):933–42.
- Ng K, May FP, Schrag D. US preventive Services task force recommendations for colorectal cancer screening: Forty-five is the new fifty. *JAMA* 2021;325(19):1943–5.
- O’Sullivan DE, Sutherland RL, Town S, et al. Risk factors for early-onset colorectal cancer: A systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2022;20(6):1229–40.e5.
- Gu WJ, Pei JP, Lyu J, et al. The burden of early-onset colorectal cancer and its risk factors from 1990 to 2019: A systematic analysis for the global burden of disease study 2019. *Cancers (Basel)* 2022;14.
- Eng C, Jacome AA, Agarwal R, et al. A comprehensive framework for early-onset colorectal cancer research. *Lancet Oncol* 2022;23(3):e116–28.
- Phipps AI, Chan AT, Ogino S. Anatomic subsite of primary colorectal cancer and subsequent risk and distribution of second cancers. *Cancer* 2013;119(17):3140–7.
- Wang L, Lo CH, He X, et al. Risk factor profiles differ for cancers of different regions of the colorectum. *Gastroenterology* 2020;159(1):241–56.e13.
- Jess P, Hansen IO, Gamborg M, et al. A nationwide Danish cohort study challenging the categorisation into right-sided and left-sided colon cancer. *BMJ Open* 2013;3(5):e002608.
- Testa U, Pelosi E, Castelli G. Colorectal cancer: Genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells. *Med Sci (Basel)* 2018;6(2):31.

51. Mima K, Kosumi K, Baba Y, et al. The microbiome, genetics, and gastrointestinal neoplasms: The evolving field of molecular pathological epidemiology to analyze the tumor-immune-microbiome interaction. *Hum Genet* 2021;140(5):725–46.
52. Ternes D, Karta J, Tsenkova M, et al. Microbiome in colorectal cancer: How to get from meta-omics to mechanism? *Trends Microbiol* 2020; 28(5):401–23.
53. Purcell RV, Visnovska M, Biggs PJ, et al. Distinct gut microbiome patterns associate with consensus molecular subtypes of colorectal cancer. *Sci Rep* 2017;7(1):11590.
54. Nguyen LH, Ma W, Wang DD, et al. Association between sulfur-metabolizing bacterial communities in stool and risk of distal colorectal cancer in men. *Gastroenterology* 2020;158(5):1313–25.
55. Mehta RS, Song M, Nishihara R, et al. Dietary patterns and risk of colorectal cancer: Analysis by tumor location and molecular subtypes. *Gastroenterology* 2017;152(5):S142–3.
56. James KR, Gomes T, Elmentaite R, et al. Distinct microbial and immune niches of the human colon. *Nat Immunol* 2020;21(3):343–53.
57. Bai J, Chen H, Bai X. Relationship between microsatellite status and immune microenvironment of colorectal cancer and its application to diagnosis and treatment. *J Clin Lab Anal* 2021;35(6):e23810.
58. Carr PR, Alwers E, Bienert S, et al. Lifestyle factors and risk of sporadic colorectal cancer by microsatellite instability status: A systematic review and meta-analyses. *Ann Oncol* 2018;29(4):825–34.
59. Niv Y. Microsatellite instability and MLH1 promoter hypermethylation in colorectal cancer. *World J Gastroenterol* 2007;13(12):1867–9.
60. Landau MA, Zhu B, Akwuole FN, et al. Site-specific differences in colonic adenocarcinoma: KRAS mutations and high tumor budding are more frequent in cecal adenocarcinoma. *Am J Surg Pathol* 2018; 42(3):351–8.
61. Holowatyj AN, Eng C, Wen W, et al. Spectrum of somatic cancer gene variations among adults with appendiceal cancer by age. *JAMA Netw Open* 2020;3(12):e2028644.
62. Tokunaga R, Xiu J, Johnston C, et al. Molecular profiling of appendiceal adenocarcinoma and comparison with right-sided and left-sided colorectal cancer. *Clin Cancer Res* 2019;25(10):3096–103.