**Supplementary Methods**

Microsatellite instability (MSI) Status

Tumor MSI status was analyzed using polymerase chain reaction (PCR) assays by all studies with the exception of NSHDS and EPIC\_Sweden. Study-specific marker panels for PCR-based MSI assays are provided in **Supplementary Table S2**. CCFR, CPS-II, HPFS, MCCS, NFCCR, and NHS used the Bethesda Consensus Panel (1) with minor variability (2-6). In DACHS (7), a mononucleotide marker panel, which has high concordance with the Bethesda Consensus Panel for the detection of MSI-high status, was used. In DALS (8), a panel of two mononucleotide and ten tetranucleotide repeats was used, and that panel has shown high concordance with the Bethesda Panel for MSI-high detection (9). For these PCR-based methods, high-level MSI (MSI-high) was defined as presence of instability in ≥30% of markers. EPIC\_Sweden and NSHDS (10) used immunohistochemical detection of deficiency for mismatch repair (MMR) proteins MLH1, MSH2, MSH6, and PMS2 (following protein nomenclature recommendations by an expert panel) (11). In EDRN, both immunohistochemistry (MLH1, MSH2, MSH6, and PMS2) and PCR (seven-marker panel) were performed to determine MMR status (12). Because studies have shown high concordance of the PCR-based test (for the detection of high-level MSI) and immunohistochemistry-based test (for the detection of MMR protein loss) (13), tumors with MMR protein loss were included as MSI-high tumors. In TCGA, we utilized MSI status designations based on calls using mSINGS (14) on sequencing data and previously published TCGA calls (15, 16). Tumors with valid data by MSI (or MMR) testing that did not show MSI-high status (or MMR protein loss) were categorized as “non-MSI-high”, which included both MSI-low and microsatellite stable tumors.

CpG island methylator phenotype (CIMP) Status

CIMP status was determined by methylation assay using different panels across studies (**Supplementary Table S3**). In CCFR, CPS-II, EPIC\_Sweden, HPFS, MCCS, NHS, and NSHDS, MethyLight assay (17) was used to determine CIMP status. CPS-II, HPFS, and NHS used an eight-gene panel; CCFR and MCCS used a five-gene panel. DALS and DACHS used methylation-specific PCR assays and different panels of five CpG islands. For TCGA data, we utilized CIMP status as determined using an array-based method. We created two CIMP categories for this analysis: CIMP-high and CIMP-low/negative. The CIMP-positive status in a certain study (18) was shown to be nearly equivalent to the CIMP-high status in other studies (17, 19).

*BRAF* and *KRAS* Mutation Status

For *BRAF*, the CCFR, CPS-II, MCCS, and NFCCR evaluated *BRAF* (c.1799T>A, p.V600E) mutations using a fluorescent allele-specific PCR assay (20, 21). DACHS used both Sanger sequencing and immunohistochemical analysis of p.V600E expression to determine *BRAF* mutation status (22). DALS evaluated *BRAF* mutations using Sanger sequencing (23). EDRN tested *BRAF* c.1799T>A mutations status using primarily real-time PCR. HPFS and NHS performed PCR and pyrosequencing to identify *BRAF* codon 600 mutation (24). NSHDS and EPIC\_Sweden used real-time PCR using an allelic discrimination assay (10) to detect *BRAF* c.1799T>A mutations. In TCGA, we utilized *BRAF* codon 600 mutation data, as previously described (16).

For *KRAS*, the CCFR, CPS-II, DALS, NSHDS, and EPIC\_Sweden used Sanger sequencing to assess mutations in *KRAS* codons 12 and 13 (10, 25-27). DACHS determined *KRAS* mutation status by a single stranded conformational polymorphism technique or by Sanger sequencing (28). EDRN tested for *KRAS* codons 12, 13, and 61 primarily using Sanger sequencing. MCCS used real-time PCR with high resolution melting analysis, followed by Sanger sequencing for positive tumors to identify *KRAS* codons 12 and 13 mutations (29). HPFS and NHS performed PCR and pyrosequencing to identify *KRAS* mutations in codons 12, 13, 61, and 146 (30). In TCGA, we utilized *KRAS* mutation data for codons 12, 13, 59, 61, 117, 146, as previously described (16). In this study, for data harmonization, *KRAS* mutation positivity was defined as the presence of any mutation in *KRAS* codon 12 or 13.

**References**

1. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58:5248-57.

2. Lindor NM, Burgart LJ, Leontovich O, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. J Clin Oncol 2002;20:1043-8.

3. Cicek MS, Lindor NM, Gallinger S, et al. Quality assessment and correlation of microsatellite instability and immunohistochemical markers among population- and clinic-based colorectal tumors results from the Colon Cancer Family Registry. J Mol Diagn 2011;13:271-81.

4. Woods MO, Hyde AJ, Curtis FK, et al. High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. Clin Cancer Res 2005;11:6853-61.

5. Ogino S, Brahmandam M, Cantor M, et al. Distinct molecular features of colorectal carcinoma with signet ring cell component and colorectal carcinoma with mucinous component. Mod Pathol 2006;19:59-68.

6. Buchanan DD, Clendenning M, Rosty C, et al. Tumor testing to identify lynch syndrome in two Australian colorectal cancer cohorts. J Gastroenterol Hepatol 2017;32:427-438.

7. Hoffmeister M, Bläker H, Kloor M, et al. Body mass index and microsatellite instability in colorectal cancer: a population-based study. Cancer Epidemiol Biomarkers Prev 2013;22:2303-11.

8. Samowitz WS, Slattery ML. Regional reproducibility of microsatellite instability in sporadic colorectal cancer. Genes Chromosomes Cancer 1999;26:106-14.

9. Slattery ML, Curtin K, Anderson K, et al. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. J Natl Cancer Inst 2000;92:1831-6.

10. Dahlin AM, Palmqvist R, Henriksson ML, et al. The role of the CpG island methylator phenotype in colorectal cancer prognosis depends on microsatellite instability screening status. Clin Cancer Res 2010;16:1845-55.

11. Fujiyoshi K, Bruford EA, Mroz P, et al. Opinion: Standardizing gene product nomenclature-a call to action. Proc Natl Acad Sci U S A 2021;118.

12. Shia J. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. J Mol Diagn 2008;10:293-300.

13. Sepulveda AR, Hamilton SR, Allegra CJ, et al. Molecular Biomarkers for the Evaluation of Colorectal Cancer: Guideline From the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and American Society of Clinical Oncology. J Mol Diagn 2017;19:187-225.

14. Salipante SJ, Scroggins SM, Hampel HL, et al. Microsatellite instability detection by next generation sequencing. Clin Chem 2014;60:1192-9.

15. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330-7.

16. Grasso CS, Giannakis M, Wells DK, et al. Genetic Mechanisms of Immune Evasion in Colorectal Cancer. Cancer Discov 2018;8:730-749.

17. Ogino S, Kawasaki T, Brahmandam M, et al. Precision and performance characteristics of bisulfite conversion and real-time PCR (MethyLight) for quantitative DNA methylation analysis. J Mol Diagn 2006;8:209-17.

18. Weisenberger DJ, Siegmund KD, Campan M, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787-93.

19. Hinoue T, Weisenberger DJ, Lange CP, et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res 2012;22:271-82.

20. Buchanan DD, Sweet K, Drini M, et al. Risk factors for colorectal cancer in patients with multiple serrated polyps: a cross-sectional case series from genetics clinics. PLoS One 2010;5:e11636.

21. Loughrey MB, Waring PM, Tan A, et al. Incorporation of somatic BRAF mutation testing into an algorithm for the investigation of hereditary non-polyposis colorectal cancer. Fam Cancer 2007;6:301-10.

22. Bläker H, Alwers E, Arnold A, et al. The Association Between Mutations in BRAF and Colorectal Cancer-Specific Survival Depends on Microsatellite Status and Tumor Stage. Clin Gastroenterol Hepatol 2019;17:455-462.e6.

23. Samowitz WS, Albertsen H, Herrick J, et al. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. Gastroenterology 2005;129:837-45.

24. 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;61:847-54.

25. Alsop K, Mead L, Smith LD, et al. Low somatic K-ras mutation frequency in colorectal cancer diagnosed under the age of 45 years. Eur J Cancer 2006;42:1357-61.

26. Samowitz WS, Curtin K, Schaffer D, et al. Relationship of Ki-ras mutations in colon cancers to tumor location, stage, and survival: a population-based study. Cancer Epidemiol Biomarkers Prev 2000;9:1193-7.

27. Myte R, Gylling B, Haggstrom J, et al. One-carbon metabolism biomarkers and genetic variants in relation to colorectal cancer risk by KRAS and BRAF mutation status. PLoS One 2018;13:e0196233.

28. Jia M, Jansen L, Walter V, et al. No association of CpG island methylator phenotype and colorectal cancer survival: population-based study. Br J Cancer 2016;115:1359-1366.

29. Rosty C, Young JP, Walsh MD, et al. Colorectal carcinomas with KRAS mutation are associated with distinctive morphological and molecular features. Mod Pathol 2013;26:825-34.

30. Imamura Y, Lochhead P, Yamauchi M, et al. Analyses of clinicopathological, molecular, and prognostic associations of KRAS codon 61 and codon 146 mutations in colorectal cancer: cohort study and literature review. Mol Cancer 2014;13:135.



Supplementary Figure 1. Prevalence of MSI-High Status, CIMP-High Status, *BRAF* Mutations, and *KRAS* Mutations along Sublocations by Sexes.

Abbreviations: CIMP, CpG island methylator phenotype; MSI, microsatellite instability.



Supplementary Figure 2. Prevalence of MSI-High Status, CIMP-High Status, *BRAF* Mutations, and *KRAS* Mutations along Sublocations by Family History of Colorectal Cancer.

Abbreviations: CIMP, CpG island methylator phenotype; MSI, microsatellite instability.

Supplementary Table 1. Patients and Molecular Characteristics of Colorectal Cancer According to Participating Studies

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Patients characteristics\* | N | CCFR | CPS II | DACHS | DALS | EDRN | EPIC\_Sweden | HPFS | MCCS | NFCCR | NHS | NSHDS | TCGA |
| All cases | 14004 | 6491 | 856 | 2007 | 1071 | 218 | 137 | 607 | 482 | 547 | 784 | 311 | 493 |
| Sex |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Female | 6712 (48%) | 3166 (49%) | 421 (49%) | 842 (42%) | 483 (45%) | 85 (39%) | 58 (42%) | 0 (0%) | 224 (46%) | 216 (39%) | 784 (100%) | 196 (63%) | 237 (48%) |
|  Male | 7292 (52%) | 3325 (51%) | 435 (51%) | 1165 (58%) | 588 (55%) | 133 (61%) | 79 (58%) | 607 (100%) | 258 (54%) | 331 (61%) | 0 (0%) | 115 (37%) | 256 (52%) |
| Mean age ± SD | 62·7±12·5 | 54.2±12.3 | 74·1±6·5 | 68·8±10·8 | 65·4±9·5 | 62·6±12·5 | 64·0±7·7 | 70·5±8·8 | 67·8±8·4 | 60·0±9·3 | 67·0±8·5 | 64·6±8·2 | 65·8±12·7 |
| Family history of CRC |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Absent | 9421 (81%) | 4192 (66%) | 658 (85%) | 1700 (86%) | 898 (84%) | 189 (87%) | N/A | 481 (79%) | 400 (83%) | 454 (90%) | 646 (84%) | 297 (96%) | 358 (86%) |
|  Present | 2,167 (19%) | 2190 (34%) | 114 (15%) | 269 (14%) | 173 (16%) | 28 (13%) | N/A | 126 (21%) | 82 (17%) | 51 (10%) | 119 (16%) | 11 (4%) | 59 (14%) |
| Primary tumor location |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Cecum | 2142 (15%) | 922 (14%) | 174 (20%) | 228 (11%) | 257 (24%) | 32 (15%) | 20 (15%) | 123 (20%) | 58 (12%) | 90 (16%) | 114 (15%) | 43 (14%) | 81 (16%) |
|  Ascending colon | 1962 (14%) | 790 (12%) | 137 (16%) | 287 (14%) | 154 (14%) | 30 (14%) | 15 (11%) | 82 (14%) | 57 (12%) | 50 (9%) | 226 (29%) | 37 (12%) | 97 (20%) |
|  Transverse colon | 1390 (10%) | 611 (9%) | 127 (15%) | 199 (10%) | 148 (14%) | 12 (6%) | 12 (9%) | 57 (9%) | 50 (10%) | 56 (10%) | 60 (8%) | 28 (9%) | 30 (6%) |
|  Descending colon | 1039 (7.4%) | 481 (7%) | 62 (7%) | 138 (7%) | 121 (11%) | 20 (9%) | 6 (4%) | 36 (6%) | 24 (5%) | 44 (8%) | 65 (8%) | 22 (7%) | 20 (4%) |
|  Sigmoid colon | 3310 (24%) | 1494 (23%) | 179 (21%) | 415 (21%) | 391 (37%) | 43 (20%) | 35 (26%) | 156 (26%) | 113 (23%) | 129 (24%) | 159 (20%) | 87 (28%) | 109 (22%) |
|  Rectum | 4161 (30%) | 2193 (34%) | 177 (21%) | 740 (37%) | 0 (0%) | 81 (37%) | 49 (36%) | 153 (25%) | 180 (37%) | 178 (33%) | 160 (20%) | 94 (30%) | 156 (32%) |
| AJCC disease stage |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  I | 3279 (25%) | 1465 (24%) | 372 (43%) | 369 (18%) | 359 (35%) | 88 (44%) | N/A | 153 (29%) | 57 (21%) | 93 (18%) | 177 (24%) | 62 (21%) | 84 (18%) |
|  II and III | 8228 (63%) | 3971 (66%) | 440 (51%) | 1352 (67%) | 578 (57%) | 88 (44%) | N/A | 302 (58%) | 164 (61%) | 368 (72%) | 459 (62%) | 190 (63%) | 316 (67%) |
|  IV | 1457 (11%) | 622 (10%) | 44 (5%) | 285 (14%) | 86 (8%) | 23 (12%) | N/A | 70 (13%) | 48 (18%) | 51 (10%) | 105 (14%) | 49 (16%) | 74 (16%) |
| MSI status |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Non-MSI-high | 10903 (83%) | 4895 (80%) | 544 (82%) | 1625 (90%) | 884 (84%) | 169 (83%) | 107 (84%) | 515 (88%) | 419 (87%) | 480 (89%) | 600 (80%) | 244 (85%) | 421 (87%) |
|  MSI-high | 2193 (17%) | 1212 (20%) | 120 (18%) | 187 (10%) | 171 (16%) | 35 (17%) | 20 (16%) | 67 (12%) | 62 (13%) | 60 (11%) | 153 (20%) | 43 (15%) | 63 (13%) |
| CIMP status |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Low/negative | 7926 (84%) | 2842 (87%) | 583 (75%) | 1679 (85%) | 697 (83%) | N/A | 99 (80%) | 450 (83%) | 401 (89%) | N/A | 562 (74%) | 227 (80%) | 386 (80%) |
|  High | 1561 (16%) | 417 (13%) | 193 (25%) | 305 (15%) | 139 (17%) | N/A | 24 (20%) | 90 (17%) | 48 (11%) | N/A | 193 (26%) | 56 (20%) | 96 (20%) |
| *BRAF*  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Wild-type | 10957 (89%) | 5087 (90%) | 583 (85%) | 1697 (93%) | 792 (90%) | 71 (81%) | 92 (75%) | 543 (92%) | 385 (86%) | 437 (90%) | 603 (79%) | 222 (78%) | 445 (91%) |
|  Mutant | 1396 (11%) | 592 (10%) | 104 (15%) | 135 (7%) | 90 (10%) | 17 (19%) | 31 (25%) | 48 (8%) | 61 (14%) | 50 (10%) | 159 (21%) | 64 (22%) | 45 (9%) |
| *KRAS*  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  Wild-type | 6710 (67%) | 2627 (67%) | 385 (68%) | 1247 (68%) | 675 (69%) | 56 (63%) | 96 (77%) | 333 (57%) | 316 (71%) | N/A | 489 (64%) | 214 (74%) | 272 (55%) |
|  Mutant | 3369 (33%) | 1270 (33%) | 178 (32%) | 590 (32%) | 310 (31%) | 33 (37%) | 29 (23%) | 256 (43%) | 132 (29%) | N/A | 276 (36%) | 76 (26%) | 219 (45%) |

\* Percentage indicates the proportion of patients with a specific patient and molecular characteristic among all patients or in strata of the studies.

Abbreviations: AJCC, American Joint Committee on Cancer; CCFR, Colon Cancer Family Registry; CIMP, CpG island methylator phenotype; CPS-II, Cancer Prevention Study II; DACHS, Darmkrebs: Chancen der Verhütung durch Screening Study; DALS, Diet, Activity and Lifestyle Study; EDRN, Early Detection Research Network; EPIC\_Sweden, European Prospective Investigation into Cancer\_Sweden; HPFS, Health Professionals Follow-up Study; MCCS, Melbourne Collaborative Cohort Study; MSI, microsatellite instability; N/A, not available; NFCCR, Newfoundland Familial Colorectal Cancer Registries; NHS, Nurses’ Health Study; NSHDS, Northern Sweden Health and Disease Study; SD, standard deviation; TCGA, The Cancer Genome Atlas.

Supplementary Table 2. Study-Specific Markers Used to Assess Microsatellite Instability (MSI) and Definition of MSI Status

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Study**† | **Mononucleotide markers** | **Dinucleotide markers** | **Other markers** | **Threshold for interpretability** | **Definitions** |
| CCFR | BAT25, BAT26, BAT40, BAT34C4 | D5S346, D17S250, ACTC, D18S55, D10S197 | MYCL | ≥4 interpretable markers | \* MSI-high if ≥30% markers showed instability  |
| CPS-II | BAT25, BAT26, BAT40, BAT34C4 | ACTC, D10S197, D17S250, D18S55, D5S346 | MYCL | ≥5 interpretable markers (unless 4 markers were unstable) | \* MSI-high if ≥30% markers showed instability  |
| DACHS | BAT25, BAT26,CAT25 | N/A | N/A | All 3 markersinterpretable | \* MSI-high if >1 marker showed instability |
| DALS | BAT26, TGFBRII | N/A | UT2127, UT2271, UT5144, UT868, UT5013, UT909, UT1205, UT5658, UT269, UT574 | ≥6/10 markers be interpretable from tetranucleotide repeat panel  | \* MSI: Instability in BAT26, *TGFBRII*, or 10 tetranucleotide marker panel.  - 10 marker panel: ≥30% unstable repeats. |
| EDRN | BAT25, BAT26, CAT25 | D5S346, D17S250, D2S123 | - | - | - |
| HPFS | BAT25, BAT26, BAT40 | D18S55, D18S56, D18S67, D18S487, D2S123, D5S346, D17S250 | - | >7 interpretable markers | \* MSI-high if ≥30% markers showed instability  |
| MCCS | BAT25, BAT26, BAT40, BAT34C4 | D5S346, D17S250, ACTC, D18S55, D10S197 | MYCL | ≥4 interpretable markers | \* MSI-high if ≥30% markers showed instability  |
| NFCCR | BAT25, BAT26, BAT40, BAT34C4 | D5S346, D17S250, ACTC, D18S55, D10S197. | MYCL | ≥4 interpretable markers | \* MSI-high if ≥30% markers showed instability  |
| NHS | BAT25, BAT26, BAT40 | D18S55, D18S56, D18S67, D18S487, D2S123, D5S346, D17S250 | - | >7 interpretable markers | \* MSI-high if ≥30% markers showed instability  |

† EPIC\_Sweden and NSHDS, and an EDRN subset used immunohistochemical detection of deficiency for mismatch repair gene proteins MLH1, MSH2, MSH6, and PMS2, and not PCR based assessment of microsatellite status.

Abbreviations: CCFR, Colon Cancer Family Registry; CPS-II, Cancer Prevention Study II; DACHS, Darmkrebs: Chancen der Verhütung durch Screening Study; DALS, Diet, Activity and Lifestyle Study; EDRN, Early Detection Research Network; HPFS, Health Professionals Follow-up Study; MCCS, Melbourne Collaborative Cohort Study; MSI, microsatellite instability, NFCCR, Newfoundland Familial Colorectal Cancer Registries; NHS, Nurses’ Health Study.

Supplementary Table 3. Study-Specific Panels Used to Assess CpG Island Methylator Phenotype (CIMP) Status

|  |  |  |  |
| --- | --- | --- | --- |
| **Study**† | **Panel genes** | **Marker positive definition** | **Definition for CIMP-high (vs. CIMP-low/negative)** |
| CCFR | *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3*, *SOCS1* | PMR >10 | \* CIMP-high: ≥3/5 methylated markers |
| CPS-II | *CDKN2A*, *MLH1*, *CACNA1G*, *NEUROG1*, *RUNX3*, *SOCS1*, *IGF2*, *CRABP1* | PMR >4 (>6 for *CRABP1*, *IGF2*) | \* CIMP-high: ≥5/8 methylated markers |
| DACHS | *MGMT*, *MLH1*, MINT1, MINT2,MINT31 |  | \* CIMP-high: ≥3/5 methylated markers |
| DALS | MINT1, MINT2, MINT31, CDKN2A9, hMLH1 |  | \* CIMP-high: ≥3/5 methylated markers |
| EPIC\_Sweden | *CDKN2A*, *MLH1*, *CACNA1G*, *NEUROG1*, *RUNX3*, *SOCS1*, *IGF2*, *CRABP1* | PMR >10 | \* CIMP-high: ≥5/8 methylated markers |
| HPFS | *CDKN2A*, *MLH1*, *CACNA1G*, *NEUROG1*, *RUNX3*, *SOCS1*, *IGF2*, *CRABP1* | PMR >4 (>6 for *CRABP1*, *IGF2*) | \* CIMP-high: ≥5/8 methylated markers |
| MCCS | *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3*, *SOCS1* | PMR >10 | \* CIMP-high: ≥3/5 methylated markers |
| NHS | *CDKN2A*, *MLH1*, *CACNA1G*, *NEUROG1*, *RUNX3*, *SOCS1*, *IGF2*, *CRABP1* | PMR >4 (>6 for *CRABP1*, *IGF2*) | \* CIMP-high: ≥5/8 methylated markers |
| NSHDS | *CDKN2A*, *MLH1*, *CACNA1G*, *NEUROG1*, *RUNX3*, *SOCS1*, *IGF2*, *CRABP1* | PMR >10 | \* CIMP-high: ≥5/8 methylated markers |

† EDRN and NFCCR did not have available CIMP data.

Abbreviations: CCFR, Colon Cancer Family Registry; CIMP, CpG island methylator phenotype; CPS-II, Cancer Prevention Study II; DACHS, Darmkrebs: Chancen der Verhütung durch Screening Study; DALS, Diet, Activity and Lifestyle Study; EPIC\_Sweden, European Prospective Investigation into Cancer\_Sweden; HPFS, Health Professionals Follow-up Study; MCCS, Melbourne Collaborative Cohort Study; NHS, Nurses’ Health Study; NSHDS, Northern Sweden Health and Disease Study; PMR, percentage of methylated reference.

Supplementary Table 4. Molecular Characteristics of Early-onset and Later-onset Colorectal Cancers According to Primary Tumor Location in Strata of Age Groups

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <40** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.76 (1.54-2.02) | <0.001 |
|  Non-MSI-high | 482 (66%) | 44 (52%) | 36 (37%) | 33 (42%) | 43 (67%) | 112 (77%) | 214 (83%) |  |  |
|  MSI-high | 243 (34%) | 40 (48%) | 61 (63%) | 45 (58%) | 21 (33%) | 33 (23%) | 43 (17%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.07 (0.68-1.69) | 0.77 |
|  Low/negative | 245 (94%) | 27 (87%) | 22 (96%) | 28 (93%) | 19 (100%) | 58 (92%) | 91 (97%) |  |  |
|  High | 15 (5.8%) | 4 (13%) | 1 (4.3%) | 2 (7%) | 0 (0%) | 5 (7.9%) | 3 (3.2%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 0.80 (0.60-1.08) | 0.15 |
|  Wild-type | 629 (95%) | 74 (95%) | 82 (99%) | 64 (94%) | 55 (96%) | 134 (94%) | 220 (95%) |  |  |
|  Mutant | 31 (4.7%) | 4 (5.1%) | 1 (1.2%) | 4 (5.9%) | 2 (3.5%) | 8 (5.6%) | 12 (5.2%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.19 (0.99-1.42) | 0.065 |
|  Wild-type | 228 (69%) | 33 (79%) | 20 (59%) | 22 (67%) | 16 (53%) | 49 (69%) | 88 (73%) |  |  |
|  Mutant | 103 (31%) | 9 (21%) | 14 (41%) | 11 (33%) | 14 (47%) | 22 (31%) | 33 (27%) |  |  |
| **Age 40-49** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.96 (1.77-2.15) | <0.001 |
|  Non-MSI-high | 1809 (83%) | 178 (68%) | 132 (61%) | 104 (63%) | 144 (78%) | 457 (93%) | 794 (94%) |  |  |
|  MSI-high | 358 (17%) | 83 (32%) | 85 (39%) | 61 (37%) | 41 (22%) | 34 (6.9%) | 54 (6.4%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.74 (1.38-2.18) | <0.001 |
|  Low/negative | 813 (94%) | 106 (86%) | 71 (83%) | 68 (93%) | 77 (96%) | 180 (97%) | 311 (97%) |  |  |
|  High | 56 (6.4%) | 17 (14%) | 15 (17%) | 5 (6.8%) | 3 (3.8%) | 5 (2.7%) | 11 (3.4%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.46 (1.17-1.68) | <0.001 |
|  Wild-type | 1900 (95%) | 236 (96%) | 179 (89%) | 148 (90%) | 158 (94%) | 446 (96%) | 733 (97%) |  |  |
|  Mutant | 102 (5.1%) | 11 (4%) | 22 (11%) | 16 (10%) | 10 (6.0%) | 20 (4.3%) | 23 (3.0%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.08 (0.98-1.20) | 0.11 |
|  Wild-type | 807 (65%) | 74 (44%) | 61 (59%) | 61 (66%) | 79 (72%) | 211 (70%) | 321 (69%) |  |  |
|  Mutant | 429 (35%) | 93 (56%) | 43 (41%) | 32 (34%) | 30 (28%) | 89 (30%) | 142 (31%) |  |  |
| **Age 50-59** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.08 (1.88-2.30) | <0.001 |
|  Non-MSI-high | 2140 (86%) | 253 (73%) | 182 (64%) | 145 (71%) | 170 (86%) | 584 (95%) | 806 (96%) |  |  |
|  MSI-high | 344 (14%) | 93 (27%) | 101 (36%) | 60 (29%) | 27 (14%) | 33 (5.3%) | 30 (3.6%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.94 (1.69-2.23) | <0.001 |
|  Low/negative | 1678 (91%) | 194 (80%) | 157 (77%) | 122 (84%) | 138 (95%) | 452 (98%) | 615 (97%) |  |  |
|  High | 159 (9%) | 48 (20%) | 48 (23%) | 23 (16%) | 8 (5.5%) | 11 (2.4%) | 21 (3.3%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.53 (1.36-1.72) | <0.001 |
|  Wild-type | 2166 (92%) | 289 (89%) | 225 (85%) | 165 (85%) | 165 (91%) | 552 (96%) | 770 (96%) |  |  |
|  Mutant | 176 (8%) | 36 (11%) | 39 (15%) | 30 (15%) | 16 (8.8%) | 23 (4.0%) | 32 (4.0%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.11 (1.03-1.20) | 0.008 |
|  Wild-type | 1233 (67%) | 139 (56%) | 123 (64%) | 86 (61%) | 97 (66%) | 347 (71%) | 441 (72%) |  |  |
|  Mutant | 597 (33%) | 111 (44%) | 70 (36%) | 54 (39%) | 50 (34%) | 140 (29%) | 172 (28%) |  |  |
| **Age 60-69** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.10 (1.93-2.28) | <0.001 |
|  Non-MSI-high | 3244 (86%) | 429 (72%) | 340 (68%) | 265 (70%) | 243 (90%) | 946 (96%) | 1021 (97%) |  |  |
|  MSI-high | 541 (14%) | 165 (28%) | 162 (32%) | 113 (30%) | 27 (10%) | 39 (4.0%) | 35 (3.3%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 2.04 (1.87-2.22) | <0.001 |
|  Low/negative | 2497 (83%) | 334 (69%) | 250 (62%) | 196 (65%) | 186 (91%) | 712 (94%) | 819 (96%) |  |  |
|  High | 515 (17%) | 150 (31%) | 156 (38%) | 105 (35%) | 19 (9.3%) | 49 (6.4%) | 36 (4.2%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.95 (1.79-2.12) | <0.001 |
|  Wild-type | 3108 (86%) | 427 (76%) | 329 (69%) | 259 (73%) | 230 (91%) | 859 (94%) | 1004 (96%) |  |  |
|  Mutant | 496 (14%) | 133 (24%) | 151 (31%) | 97 (27%) | 22 (8.7%) | 56 (6.1%) | 37 (3.6%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.01 (0.96-1.07) | 0.67 |
|  Wild-type | 2141 (67%) | 283 (56%) | 295 (68%) | 224 (70%) | 154 (67%) | 574 (70%) | 611 (68%) |  |  |
|  Mutant | 1061 (33%) | 218 (44%) | 141 (32%) | 98 (30%) | 75 (33%) | 242 (30%) | 287 (32%) |  |  |
|  |  |  |  |  |  |  |  |  |  |
| **Age ≥70** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.27 (2.08-2.47) | <0.001 |
|  Non-MSI-high | 3228 (82%) | 536 (74%) | 466 (62%) | 301 (68%) | 229 (86%) | 843 (95%) | 853 (98%) |  |  |
|  MSI-high | 707 (18%) | 189 (26%) | 286 (38%) | 139 (32%) | 38 (14%) | 40 (4.5%) | 15 (1.7%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.95 (1.82-2.09) | <0.001 |
|  Low/negative | 2693 (77%) | 442 (69%) | 356 (53%) | 266 (64%) | 190 (84%) | 718 (93%) | 721 (93%) |  |  |
|  High | 816 (23%) | 203 (31%) | 312 (47%) | 151 (36%) | 36 (16%) | 58 (7.5%) | 56 (7.2%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 2.15 (1.97-2.34) | <0.001 |
|  Wild-type | 3154 (84%) | 523 (76%) | 482 (67%) | 310 (74%) | 220 (91%) | 800 (95%) | 819 (98%) |  |  |
|  Mutant | 591 (16%) | 162 (24%) | 241 (33%) | 107 (26%) | 23 (9.5%) | 38 (4.5%) | 20 (2.4%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.93 (0.88-0.98) | 0.008 |
|  Wild-type | 2301 (66%) | 345 (55%) | 458 (68%) | 306 (78%) | 152 (66%) | 546 (70%) | 494 (63%) |  |  |
|  Mutant | 1179 (34%) | 279 (45%) | 211 (32%) | 87 (22%) | 79 (34%) | 233 (30%) | 290 (37%) |  |  |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., MSI-high, CIMP-high, *BRAF* mutation, or *KRAS* mutation) and Ptrend 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).

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

Supplementary Table 5. Molecular Characteristics of Early-onset and Later-onset According to Primary Tumor Location in Strata of Family History of Colorectal Cancer

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <50 (early-onset)** |  |  |  |  |  |  |  |  |  |
|  **Family history of CRC absent** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.71 (1.54-1.89) | <0.001 |
|  Non-MSI-high | 1786 (88%) | 172 (79%) | 135 (70%) | 108 (72%) | 153 (85%) | 440 (93%) | 172 (79%) |  |  |
|  MSI-high | 255 (12%) | 47 (21%) | 59 (30%) | 41 (28%) | 28 (15%) | 32 (6.8%) | 48 (5.8%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.46 (1.16-1.83) | 0.001 |
|  Low/negative | 761 (94%) | 87 (88%) | 62 (86%) | 61 (94%) | 70 (96%) | 184 (96%) | 87 (88%) |  |  |
|  High | 49 (6.0%) | 12 (12%) | 10 (14%) | 4 (6.2%) | 3 (4.1%) | 8 (4.2%) | 12 (3.9%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.35 (1.17-1.56) | <0.001 |
|  Wild-type | 1805 (95%) | 197 (94%) | 162 (91%) | 127 (89%) | 156 (96%) | 442 (95%) | 197 (94%) |  |  |
|  Mutant | 101 (5.3%) | 17 (9%) | 12 (6%) | 17 (9%) | 16 (11%) | 7 (4.9%) | 23 (3.5%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.10 (0.99-1.22) | 0.062 |
|  Wild-type | 805 (67%) | 73 (47%) | 55 (57%) | 58 (70%) | 79 (70%) | 219 (73%) | 73 (47%) |  |  |
|  Mutant | 400 (33%) | 82 (53%) | 42 (43%) | 25 (30%) | 34 (30%) | 82 (27%) | 135 (30%) |  |  |
|  **Family history of CRC present** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.11 (1.84-2.41) | <0.001 |
|  Non-MSI-high | 463 (59%) | 46 (38%) | 32 (29%) | 27 (31%) | 31 (50%) | 117 (78%) | 210 (81%) |  |  |
|  MSI-high | 324 (41%) | 74 (62%) | 78 (71%) | 60 (69%) | 31 (50%) | 33 (22%) | 48 (19%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.96 (1.25-3.09) | 0.004 |
|  Low/negative | 273 (93%) | 44 (83%) | 30 (86%) | 33 (92%) | 23 (100%) | 47 (96%) | 96 (99%) |  |  |
|  High | 20 (6.8%) | 9 (17%) | 5 (14%) | 3 (8.3%) | 0 (0%) | 2 (4.1%) | 1 (1.0%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.12 (0.87-1.44) | 0.37 |
|  Wild-type | 679 (96%) | 107 (97%) | 95 (95%) | 80 (95%) | 51 (91%) | 127 (97%) | 219 (96%) |  |  |
|  Mutant | 30 (4.2%) | 3 (2.7%) | 5 (5.0%) | 4 (4.8%) | 5 (8.9%) | 4 (3.1%) | 9 (3.9%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.11 (0.93-1.31) | 0.24 |
|  Wild-type | 218 (64%) | 32 (62%) | 25 (64%) | 25 (60%) | 15 (63%) | 39 (60%) | 82 (69%) |  |  |
|  Mutant | 123 (36%) | 20 (38%) | 14 (36%) | 17 (40%) | 9 (38%) | 26 (40%) | 37 (31%) |  |  |
| **Age ≥50 (later-onset)** |  |  |  |  |  |  |  |  |  |
|  **Family history of CRC absent** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.27 (2.12-2.42) | <0.001 |
|  Non-MSI-high | 6566 (87%) | 964 (77%) | 726 (66%) | 519 (70%) | 467 (91%) | 1821 (96%) | 2069 (98%) |  |  |
|  MSI-high | 1019 (13%) | 281 (23%) | 366 (34%) | 218 (30%) | 45 (8.8%) | 68 (3.6%) | 41 (1.9%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.99 (1.88-2.11) | <0.001 |
|  Low/negative | 5292 (82%) | 760 (73%) | 562 (59%) | 420 (65%) | 365 (89%) | 1486 (94%) | 1699 (95%) |  |  |
|  High | 1133 (18%) | 284 (27%) | 395 (41%) | 226 (35%) | 47 (11%) | 90 (5.7%) | 91 (5.1%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.98 (1.87-2.11) | <0.001 |
|  Wild-type | 6341 (87%) | 966 (82%) | 722 (69%) | 520 (74%) | 441 (92%) | 1700 (95%) | 1992 (97%) |  |  |
|  Mutant | 919 (13%) | 215 (18%) | 329 (31%) | 181 (26%) | 40 (8.3%) | 88 (4.9%) | 66 (3.2%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.97 (0.94-1.01) | 0.20 |
|  Wild-type | 4396 (66%) | 581 (54%) | 664 (68%) | 488 (74%) | 293 (66%) | 1152 (70%) | 1218 (67%) |  |  |
|  Mutant | 2220 (34%) | 491 (46%) | 316 (32%) | 170 (26%) | 149 (34%) | 495 (30%) | 599 (33%) |  |  |
|  **Family history of CRC present** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.89 (1.73-2.06) | <0.001 |
|  Non-MSI-high | 1714 (77%) | 215 (60%) | 221 (58%) | 170 (68%) | 151 (77%) | 457 (92%) | 500 (93%) |  |  |
|  MSI-high | 502 (23%) | 141 (40%) | 158 (42%) | 81 (32%) | 44 (23%) | 42 (8%) | 36 (7%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.89 (1.68-2.14) | <0.001 |
|  Low/negative | 1331 (83%) | 180 (66%) | 172 (64%) | 149 (79%) | 129 (90%) | 330 (94%) | 371 (96%) |  |  |
|  High | 280 (17%) | 91 (34%) | 95 (36%) | 40 (21%) | 15 (10%) | 22 (6%) | 17 (4%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.71 (1.53-1.92) | <0.001 |
|  Wild-type | 1791 (87%) | 240 (72%) | 270 (77%) | 194 (82%) | 156 (89%) | 433 (95%) | 498 (97%) |  |  |
|  Mutant | 278 (13%) | 95 (28%) | 82 (23%) | 42 (18%) | 19 (11%) | 24 (5%) | 16 (3%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.04 (0.96-1.13) | 0.28 |
|  Wild-type | 1064 (67%) | 162 (63%) | 174 (65%) | 113 (66%) | 97 (67%) | 262 (72%) | 256 (67%) |  |  |
|  Mutant | 526 (33%) | 96 (37%) | 92 (35%) | 59 (34%) | 48 (33%) | 103 (28%) | 128 (33%) |  |  |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., MSI-high, CIMP-high, *BRAF* mutation, or *KRAS* mutation) and Ptrend 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) and study (i.e., cohort).

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

Supplementary Table 6. Jass Classification of Early-onset and Later-onset Colorectal Cancers According to Primary Tumor Location

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <50 (early-onset)** |  |  |  |  |  |  |  |  |  |
|  Tumor subtype§ |  |  |  |  |  |  |  |  |  |
|  Type 1 | 1 (0.2%) | 0 (0%) | 0 (0%) | 0 (0%) | 1 (1.7%) | 0 (0%) | 0 (0%) | 1.43 (0.51-3.97) | 0.49 |
|  Type 2 | 17 (2.6%) | 5 (5.4%) | 4 (7.5%) | 0 (0%) | 1 (1.7%) | 5 (3.2%) | 2 (0.8%) | 1.47 (1.14-1.91) | 0.003 |
|  Type 3 | 201 (31%) | 42 (46%) | 18 (34%) | 9 (20%) | 15 (25%) | 45 (29%) | 72 (29%) | 1.14 (1.03-1.25) | <0.001 |
|  Type 4 | 376 (57%) | 28 (30%) | 15 (28%) | 26 (59%) | 39 (65%) | 104 (66%) | 164 (65%) | 0.72 (0.65-0.79) | <0.001 |
|  Type 5 | 62 (9.4%) | 17 (18%) | 16 (30%) | 9 (20%) | 4 (6.7%) | 3 (1.9%) | 13 (5.2%) | 1.55 (1.34-1.78) | <0.001 |
| **Age ≥50 (later-onset)** |  |  |  |  |  |  |  |  |  |
|  Tumor subtype§ |  |  |  |  |  |  |  |  |  |
|  Type 1 | 511 (8.3%) | 157 (16%) | 222 (25%) | 91 (17%) | 16 (3.6%) | 23 (1.4%) | 2 (0.1%) | 1.89 (1.77-2.02) | <0.001 |
|  Type 2 | 202 (3.2%) | 45 (4.7%) | 70 (8.0%) | 40 (7.3%) | 13 (3.0%) | 24 (1.5%) | 10 (0.6%) | 1.46 (1.35-1.59) | <0.001 |
|  Type 3 | 1992 (32%) | 396 (41%) | 268 (31%) | 149 (27%) | 154 (35%) | 463 (29%) | 562 (32%) | 1.06 (1.03-1.09) | <0.001 |
|  Type 4 | 3287 (53%) | 302 (31%) | 253 (29%) | 242 (44%) | 242 (55%) | 1079 (67%) | 1169 (66%) | 0.71 (0.69-0.73) | <0.001 |
|  Type 5 | 195 (3.2%) | 60 (6.3%) | 62 (7.1%) | 25 (4.6%) | 15 (3.4%) | 18 (1.1%) | 15 (0.9%) | 1.58 (1.46-1.72) | <0.001 |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity and Ptrend 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).

§ Tumor subtypes described by Jass38 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; Type 5 = MSI-high, CIMP-low/negative, *BRAF* wild-type, *KRAS* wild-type.

Abbreviations: AJCC, American Joint Committee on Cancer; CIMP, CpG island methylator phenotype; MSI, microsatellite instability.

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

Supplementary Table 7. Molecular Characteristics of Early-onset and Later-onset According to Primary Tumor Location in Strata of CIMP Status

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <50 (early-onset)** |  |  |  |  |  |  |  |  |  |
|  **CIMP-low/negative** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.04 (1.75-2.37) | <0.001 |
|  Non-MSI-high | 863 (84%) | 93 (70%) | 56 (62%) | 56 (61%) | 74 (79%) | 215 (95%) | 369 (93%) |  |  |
|  MSI-high | 167 (16%) | 39 (30%) | 35 (38%) | 36 (39%) | 20 (21%) | 11 (4.9%) | 26 (6.6%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.30 (1.03-1.64) | 0.029 |
|  Wild-type | 1004 (97%) | 129 (98%) | 86 (93%) | 86 (95%) | 89 (95%) | 230 (97%) | 384 (97%) |  |  |
|  Mutant | 35 (3.4%) | 2 (1.5%) | 6 (6.5%) | 5 (5.5%) | 5 (5.3%) | 7 (3.0%) | 10 (2.5%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.07 (0.94-1.21) | 0.30 |
|  Wild-type | 477 (65%) | 46 (48%) | 38 (58%) | 43 (69%) | 49 (71%) | 115 (69%) | 186 (69%) |  |  |
|  Mutant | 252 (35%) | 50 (52%) | 27 (42%) | 19 (31%) | 20 (29%) | 51 (31%) | 85 (31%) |  |  |
|  **CIMP-high** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.48 (0.81-2.70) | 0.20 |
|  Non-MSI-high | 48 (70%) | 11 (55%) | 11 (69%) | 2 (33%) | 2 (67%) | 9 (90%) | 13 (93%) |  |  |
|  MSI-high | 21 (30%) | 9 (45%) | 5 (31%) | 4 (67%) | 1 (33%) | 1 (10%) | 1 (7.1%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 0.73 (0.44-1.21) | 0.23 |
|  Wild-type | 41 (60%) | 12 (63%) | 11 (69%) | 5 (71%) | 1 (33%) | 3 (30%) | 9 (69%) |  |  |
|  Mutant | 27 (40%) | 7 (37%) | 5 (31%) | 2 (29%) | 2 (67%) | 7 (70%) | 4 (31%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.42 (0.83-2.43) | 0.20 |
|  Wild-type | 39 (66%) | 11 (58%) | 8 (67%) | 3 (60%) | 2 (67%) | 6 (75%) | 9 (75%) |  |  |
|  Mutant | 20 (34%) | 8 (42%) | 4 (33%) | 2 (40%) | 1 (33%) | 2 (25%) | 3 (25%) |  |  |
| **Age ≥50 (later-onset)** |  |  |  |  |  |  |  |  |  |
|  **CIMP-low/negative** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.95 (1.80-2.13) | <0.001 |
|  Non-MSI-high | 6121 (94%) | 802 (88%) | 614 (84%) | 468 (87%) | 460 (94%) | 1765 (98%) | 2012 (98%) |  |  |
|  MSI-high | 399 (6.1%) | 108 (12%) | 113 (16%) | 70 (13%) | 30 (6.1%) | 43 (2.4%) | 35 (1.7%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.41 (1.29-1.55) | <0.001 |
|  Wild-type | 6241 (96%) | 866 (96%) | 674 (92%) | 494 (93%) | 464 (97%) | 1750 (98%) | 1993 (97%) |  |  |
|  Mutant | 249 (3.8%) | 38 (4.2%) | 56 (7.7%) | 39 (7.3%) | 16 (3.3%) | 39 (2.2%) | 61 (3.0%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.13 (1.09-1.18) | <0.001 |
|  Wild-type | 3934 (64%) | 411 (48%) | 382 (55%) | 327 (64%) | 291 (63%) | 1203 (70%) | 1320 (68%) |  |  |
|  Mutant | 2251 (36%) | 444 (52%) | 314 (45%) | 182 (36%) | 170 (37%) | 522 (30%) | 619 (32%) |  |  |
|  **CIMP-high** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.75 (1.57-1.96) | <0.001 |
|  Non-MSI-high | 632 (45%) | 157 (42%) | 173 (35%) | 90 (35%) | 35 (56%) | 80 (70%) | 97 (95%) |  |  |
|  MSI-high | 768 (55%) | 217 (58%) | 319 (65%) | 165 (65%) | 28 (44%) | 34 (30%) | 5 (5%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.44 (1.30-1.60) | <0.001 |
|  Wild-type | 604 (44%) | 148 (39%) | 176 (36%) | 101 (41%) | 31 (52%) | 61 (54%) | 87 (84%) |  |  |
|  Mutant | 782 (56%) | 227 (61%) | 310 (64%) | 148 (59%) | 29 (48%) | 52 (46%) | 16 (16%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.72 (0.64-0.80) | <0.001 |
|  Wild-type | 1082 (79%) | 269 (75%) | 408 (85%) | 220 (88%) | 53 (88%) | 77 (71%) | 55 (52%) |  |  |
|  Mutant | 282 (21%) | 91 (25%) | 72 (15%) | 29 (12%) | 7 (12%) | 32 (29%) | 51 (48%) |  |  |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., MSI-high, *BRAF* mutation, or *KRAS* mutation) and Ptrend 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).

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

Supplementary Table 8. Molecular Characteristics of Early-onset and Later-onset According to Primary Tumor Location in Strata of *BRAF* Status

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <50 (early-onset)** |  |  |  |  |  |  |  |  |  |
|  ***BRAF*-wild-type** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.92 (1.75-2.09) | <0.001 |
|  Non-MSI-high | 1875 (80%) | 189 (65%) | 135 (54%) | 105 (56%) | 146 (74%) | 478 (90%) | 822 (92%) |  |  |
|  MSI-high | 478 (20%) | 102 (35%) | 115 (46%) | 83 (44%) | 52 (26%) | 53 (10%) | 73 (8.2%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.71 (1.32-2.21) | <0.001 |
|  Low/negative | 1004 (96%) | 129 (91%) | 86 (89%) | 86 (95%) | 89 (99%) | 230 (99%) | 384 (98%) |  |  |
|  High | 41 (4%) | 12 (9%) | 11 (11%) | 5 (5.5%) | 1 (1.1%) | 3 (1.3%) | 9 (2.3%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.16 (1.06-1.26) | <0.001 |
|  Wild-type | 949 (65%) | 92 (48%) | 69 (55%) | 69 (62%) | 84 (66%) | 247 (70%) | 388 (70%) |  |  |
|  Mutant | 519 (35%) | 99 (52%) | 57 (45%) | 43 (38%) | 43 (34%) | 108 (30%) | 169 (30%) |  |  |
|  ***BRAF*-mutant** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.47 (0.94-2.29) | 0.093 |
|  Non-MSI-high | 104 (87%) | 11 (85%) | 18 (86%) | 13 (76%) | 9 (75%) | 27 (100%) | 26 (90%) |  |  |
|  MSI-high | 15 (13%) | 2 (15%) | 3 (14%) | 4 (24%) | 3 (25%) | 0 (0%) | 3 (10%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.28 (0.79-2.07) | 0.32 |
|  Low/negative | 35 (56%) | 2 (22%) | 6 (55%) | 5 (71%) | 5 (71%) | 7 (50%) | 10 (71%) |  |  |
|  High | 27 (44%) | 7 (78%) | 5 (45%) | 2 (29%) | 2 (29%) | 7 (50%) | 4 (29%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.41 (0.05-2.92) | 0.38 |
|  Wild-type | 67 (97%) | 10 (100%) | 11 (100%) | 9 (100%) | 9 (100%) | 12 (92%) | 16 (94%) |  |  |
|  Mutant | 2 (2.9%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 1 (7.7%) | 1 (5.9%) |  |  |
| **Age ≥50 (later-onset)** |  |  |  |  |  |  |  |  |  |
|  ***BRAF*-wild-type** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.95 (1.83-2.08) | <0.001 |
|  Non-MSI-high | 7253 (91%) | 1011 (85%) | 763 (78%) | 551 (80%) | 537 (91%) | 2038 (97%) | 2353 (98%) |  |  |
|  MSI-high | 709 (9%) | 179 (15%) | 212 (22%) | 138 (20%) | 52 (8.8%) | 69 (3.3%) | 59 (2.4%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.66 (1.55-1.77) | <0.001 |
|  Low/negative | 6241 (91%) | 866 (85%) | 674 (79%) | 494 (83%) | 464 (94%) | 1750 (97%) | 1993 (96%) |  |  |
|  High | 604 (9%) | 148 (15%) | 176 (21%) | 101 (17%) | 31 (6.3%) | 61 (3.4%) | 87 (4.2%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.13 (1.08-1.17) | <0.001 |
|  Wild-type | 4383 (62%) | 480 (46%) | 472 (54%) | 387 (64%) | 322 (62%) | 1298 (69%) | 1424 (67%) |  |  |
|  Mutant | 2685 (38%) | 574 (54%) | 402 (46%) | 217 (36%) | 198 (38%) | 580 (31%) | 714 (33%) |  |  |
|  ***BRAF*-mutant** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.88 (1.65-2.14) | <0.001 |
|  Non-MSI-high | 498 (41%) | 91 (28%) | 129 (31%) | 84 (38%) | 33 (56%) | 83 (73%) | 78 (94%) |  |  |
|  MSI-high | 716 (59%) | 230 (72%) | 288 (69%) | 137 (62%) | 26 (44%) | 30 (27%) | 5 (6.0%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.80 (1.58-2.06) | <0.001 |
|  Low/negative | 249 (24%) | 38 (14%) | 56 (15%) | 39 (21%) | 16 (36%) | 39 (43%) | 61 (79%) |  |  |
|  High | 782 (76%) | 227 (86%) | 310 (85%) | 148 (79%) | 29 (64%) | 52 (57%) | 16 (21%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.67 (0.53-0.85) | 0.001 |
|  Wild-type | 1024 (95%) | 256 (95%) | 374 (98%) | 181 (95%) | 54 (98%) | 96 (94%) | 63 (85%) |  |  |
|  Mutant | 49 (4.6%) | 13 (4.8%) | 8 (2.1%) | 10 (5.2%) | 1 (1.8%) | 6 (5.9%) | 11 (15%) |  |  |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., MSI-high, CIMP-high, *KRAS* mutation)and Ptrend 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).

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

Supplementary Table 9. Molecular Characteristics of Early-onset and Later-onset According to Primary Tumor Location in Strata of *KRAS* Status

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <50 (early-onset)** |  |  |  |  |  |  |  |  |  |
|  ***KRAS*-wild-type** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.12 (1.81-2.50) | <0.001 |
|  Non-MSI-high | 813 (85%) | 68 (67%) | 42 (55%) | 48 (66%) | 78 (86%) | 224 (95%) | 353 (94%) |  |  |
|  MSI-high | 139 (15%) | 33 (33%) | 34 (45%) | 25 (34%) | 13 (14%) | 13 (5%) | 21 (6%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.45 (1.11-1.90) | 0.007 |
|  Low/negative | 477 (92%) | 46 (81%) | 38 (83%) | 43 (93%) | 49 (96%) | 115 (95%) | 186 (95%) |  |  |
|  High | 39 (8%) | 11 (19%) | 8 (17%) | 3 (7%) | 2 (4%) | 6 (5%) | 9 (5%) |  |  |
|  *BRAF* |  |  |  |  |  |  |  | 1.49 (1.23-1.80) | <0.001 |
|  Wild-type | 949 (93%) | 92 (90%) | 69 (86%) | 69 (88%) | 84 (90%) | 247 (95%) | 388 (96%) |  |  |
|  Mutant | 67 (7%) | 10 (10%) | 11 (14%) | 9 (12%) | 9 (10%) | 12 (5%) | 16 (4%) |  |  |
|  ***KRAS*-mutant** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.68 (1.35-2.07) | 0.093 |
|  Non-MSI-high | 408 (84%) | 82 (85%) | 37 (67%) | 21 (57%) | 30 (75%) | 90 (94%) | 148 (92%) |  |  |
|  MSI-high | 78 (16%) | 15 (15%) | 18 (33%) | 16 (43%) | 10 (25%) | 6 (6.3%) | 13 (8.1%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.99 (1.20-3.31) | 0.008 |
|  Low/negative | 252 (93%) | 50 (86%) | 27 (87%) | 19 (90%) | 20 (95%) | 51 (96%) | 85 (97%) |  |  |
|  High | 20 (7%) | 8 (14%) | 4 (13%) | 2 (10%) | 1 (4.8%) | 2 (3.8%) | 3 (3.4%) |  |  |
|  *BRAF* |  |  |  |  |  |  |  | 0.49 (0.08-3.01) | 0.44 |
|  Wild-type | 519 (100%) | 99 (100%) | 57 (100%) | 43 (100%) | 43 (100%) | 108 (99%) | 169 (99%) |  |  |
|  Mutant | 2 (0.4%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 1 (0.9%) | 1 (0.6%) |  |  |
| **Age ≥50 (later-onset)** |  |  |  |  |  |  |  |  |  |
|  ***KRAS*-wild-type** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.56 (1.38-1.77) | <0.001 |
|  Non-MSI-high | 4378 (81%) | 455 (61%) | 438 (52%) | 381 (66%) | 336 (87%) | 1351 (96%) | 1417 (98%) |  |  |
|  MSI-high | 1017 (19%) | 288 (39%) | 401 (48%) | 198 (34%) | 52 (13%) | 53 (3.8%) | 25 (1.7%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 2.31 (2.16-2.47) | <0.001 |
|  Low/negative | 3934 (78%) | 411 (60%) | 382 (48%) | 327 (60%) | 291 (85%) | 1203 (94%) | 1320 (96%) |  |  |
|  High | 1082 (22%) | 269 (40%) | 408 (52%) | 220 (40%) | 53 (15%) | 77 (6.0%) | 55 (4.0%) |  |  |
|  *BRAF* |  |  |  |  |  |  |  | 2.09 (1.96-2.22) | <0.001 |
|  Wild-type | 4383 (81%) | 480 (65%) | 472 (56%) | 387 (68%) | 322 (86%) | 1298 (93%) | 1424 (96%) |  |  |
|  Mutant | 1024 (19%) | 256 (35%) | 374 (44%) | 181 (32%) | 54 (14%) | 96 (6.9%) | 63 (4.2%) |  |  |
|  ***KRAS*-mutant** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.31 (2.16-2.47) | <0.001 |
|  Non-MSI-high | 2553 (94%) | 542 (92%) | 365 (90%) | 199 (85%) | 186 (95%) | 567 (96%) | 694 (98%) |  |  |
|  MSI-high | 167 (6.1%) | 47 (8%) | 39 (10%) | 35 (15%) | 10 (5.1%) | 22 (3.7%) | 14 (2.0%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.34 (1.22-1.48) | <0.001 |
|  Low/negative | 2251 (89%) | 444 (83%) | 314 (81%) | 182 (86%) | 170 (96%) | 522 (94%) | 619 (92%) |  |  |
|  High | 282 (11%) | 91 (17%) | 72 (19%) | 29 (14%) | 7 (4.0%) | 32 (5.8%) | 51 (7.6%) |  |  |
|  *BRAF* |  |  |  |  |  |  |  | 1.24 (1.00-1.54) | 0.050 |
|  Wild-type | 2685 (98%) | 574 (98%) | 402 (98%) | 217 (96%) | 198 (99%) | 580 (99%) | 714 (98%) |  |  |
|  Mutant | 49 (1.8%) | 13 (2.2%) | 8 (2.0%) | 10 (4.4%) | 1 (0.5%) | 6 (1.0%) | 11 (1.5%) |  |  |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., MSI-high, CIMP-high, or *BRAF* mutation)and Ptrend 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).

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

Supplementary Table 10. Molecular Characteristics of Early-onset and Later-onset MSI-high Colorectal Cancers According to Primary Tumor Location in Strata of Family History of Colorectal Cancer

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature of MSI-high cancer\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <50 (early-onset)** |  |  |  |  |  |  |  |  |  |
|  **Family history of CRC absent** |  |  |  |  |  |  |  |  |  |
|  CIMP status |  |  |  |  |  |  |  | 0.82 (0.42-1.57) | 0.54 |
|  Low/negative | 77 (91%) | 18 (90%) | 15 (88%) | 16 (94%) | 9 (90%) | 6 (86%) | 13 (93%) |  |  |
|  High | 8 (9.4%) | 2 (10%) | 2 (12%) | 1 (5.9%) | 1 (10%) | 1 (14%) | 1 (7.1%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.19 (0.67-2.11) | 0.56 |
|  Wild-type | 209 (96%) | 42 (98%) | 49 (98%) | 31 (91%) | 24 (92%) | 28 (100%) | 35 (97%) |  |  |
|  Mutant | 8 (3.7%) | 1 (2.3%) | 1 (2.0%) | 3 (8.8%) | 2 (7.7%) | 0 (0%) | 1 (2.8%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.06 (0.79-1.42) | 0.71 |
|  Wild-type | 78 (66%) | 15 (63%) | 18 (67%) | 12 (67%) | 9 (56%) | 10 (77%) | 14 (70%) |  |  |
|  Mutant | 40 (34%) | 9 (38%) | 9 (33%) | 6 (33%) | 7 (44%) | 3 (23%) | 6 (30%) |  |  |
|  **Family history of CRC present** |  |  |  |  |  |  |  |  |  |
|  CIMP status |  |  |  |  |  |  |  | 2.31 (0.72-7.41) | 0.15 |
|  Low/negative | 87 (87%) | 21 (75%) | 19 (86%) | 19 (86%) | 10 (100%) | 5 (100%) | 13 (100%) |  |  |
|  High | 13 (13%) | 7 (25%) | 3 (14%) | 3 (14%) | 0 (0%) | 0 (0%) | 0 (0%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 0.95 (0.55-1.65) | 0.86 |
|  Wild-type | 256 (97%) | 59 (98%) | 62 (97%) | 48 (98%) | 26 (96%) | 24 (100%) | 37 (95%) |  |  |
|  Mutant | 7 (2.7%) | 1 (1.7%) | 2 (3.1%) | 1 (2.0%) | 1 (3.7%) | 0 (0%) | 2 (5.1%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.02 (0.71-1.45) | 0.92 |
|  Wild-type | 61 (62%) | 18 (75%) | 16 (67%) | 13 (57%) | 4 (57%) | 3 (50%) | 7 (50%) |  |  |
|  Mutant | 37 (38%) | 6 (25%) | 8 (33%) | 10 (43%) | 3 (43%) | 3 (50%) | 7 (50%) |  |  |
| **Age ≥50 (later-onset)** |  |  |  |  |  |  |  |  |  |
|  **Family history of CRC absent** |  |  |  |  |  |  |  |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.69 (1.43-2.01) | <0.001 |
|  Low/negative | 255 (31%) | 70 (32%) | 70 (23%) | 42 (24%) | 16 (48%) | 34 (59%) | 23 (82%) |  |  |
|  High | 567 (69%) | 146 (68%) | 240 (77%) | 135 (76%) | 17 (52%) | 24 (41%) | 5 (18%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.50 (1.29-1.75) | <0.001 |
|  Wild-type | 435 (46%) | 115 (44%) | 132 (38%) | 88 (45%) | 28 (67%) | 43 (65%) | 29 (85%) |  |  |
|  Mutant | 506 (54%) | 146 (56%) | 211 (62%) | 107 (55%) | 14 (33%) | 23 (35%) | 5 (15%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.66 (0.54-0.80) | <0.001 |
|  Wild-type | 735 (86%) | 198 (85%) | 295 (91%) | 152 (86%) | 31 (86%) | 43 (74%) | 16 (59%) |  |  |
|  Mutant | 117 (14%) | 34 (15%) | 28 (9%) | 24 (14%) | 5 (14%) | 15 (26%) | 11 (41%) |  |  |
|  **Family history of CRC present** |  |  |  |  |  |  |  |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.49 (1.10-2.01) | 0.009 |
|  Low/negative | 132 (45%) | 33 (37%) | 40 (39%) | 27 (54%) | 13 (54%) | 9 (53%) | 10 (100%) |  |  |
|  High | 162 (55%) | 57 (63%) | 63 (61%) | 23 (46%) | 11 (46%) | 8 (47%) | 0 (0%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.91 (1.45-2.51) | <0.001 |
|  Wild-type | 245 (58%) | 56 (44%) | 68 (51%) | 45 (65%) | 23 (68%) | 25 (81%) | 28 (100%) |  |  |
|  Mutant | 176 (42%) | 70 (56%) | 65 (49%) | 24 (35%) | 11 (32%) | 6 (19%) | 0 (0%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.68 (0.50-0.93) | 0.015 |
|  Wild-type | 241 (84%) | 78 (90%) | 87 (89%) | 40 (78%) | 20 (80%) | 9 (60%) | 7 (70%) |  |  |
|  Mutant | 45 (16%) | 9 (10%) | 11 (11%) | 11 (22%) | 5 (20%) | 6 (40%) | 3 (30%) |  |  |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., CIMP-high, *BRAF* mutation, or *KRAS* mutation) and Ptrend 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) and study (i.e., cohort).

Abbreviations: CI, confidence interval; CIMP, CpG island methylator phenotype; CRC, colorectal cancer; MSI, microsatellite instability; OR, odds ratio.

Supplementary Table 11. Molecular Characteristics of Early-onset and Later-onset According to Primary Tumor Location in Strata of Year of Diagnosis (up to 2000 vs. After 2000).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Molecular feature\* | Total N | Cecum | Ascending colon | Transverse colon | Descending colon | Sigmoid colon | Rectum | Multivariable OR (95% CI) † | Ptrend† |
| **Age <50 (early-onset)** |  |  |  |  |  |  |  |  |  |
|  **Diagnosed up to 2000** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.78 (1.57-2.03) | <0.001 |
|  Non-MSI-high | 895 (78%) | 105 (67%) | 70 (56%) | 60 (57%) | 71 (71%) | 229 (90%) | 360 (88%) |  |  |
|  MSI-high | 255 (22%) | 51 (33%) | 56 (44%) | 46 (43%) | 29 (29%) | 26 (10%) | 47 (12%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.80 (1.39-2.33) | <0.001 |
|  Low/negative | 699 (94%) | 90 (87%) | 63 (85%) | 64 (94%) | 64 (96%) | 166 (95%) | 252 (99%) |  |  |
|  High | 43 (5.8%) | 14 (13%) | 11 (15%) | 4 (5.9%) | 3 (4.5%) | 8 (4.6%) | 3 (1.2%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.14 (1.00-1.30) | 0.008 |
|  Wild-type | 985 (94%) | 136 (96%) | 104 (91%) | 88 (93%) | 82 (92%) | 224 (93%) | 351 (97%) |  |  |
|  Mutant | 60 (5.7%) | 6 (4.2%) | 10 (8.8%) | 7 (7.4%) | 7 (7.9%) | 18 (7.4%) | 12 (3.3%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.28 (1.07-1.54) | 0.046 |
|  Wild-type | 467 (67%) | 52 (50%) | 37 (55%) | 43 (69%) | 47 (72%) | 116 (75%) | 172 (71%) |  |  |
|  Mutant | 229 (33%) | 52 (50%) | 30 (45%) | 19 (31%) | 18 (28%) | 39 (25%) | 71 (29%) |  |  |
|  **Diagnosed after 2000** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 1.95 (1.76-2.16) | <0.001 |
|  Non-MSI-high | 1396 (80%) | 117 (62%) | 98 (52%) | 77 (56%) | 116 (78%) | 340 (89%) | 648 (93%) |  |  |
|  MSI-high | 346 (20%) | 72 (38%) | 90 (48%) | 60 (44%) | 33 (22%) | 41 (11%) | 50 (7.2%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.28 (0.94-1.75) | 0.12 |
|  Low/negative | 359 (93%)  | 43 (86%) | 30 (86%) | 32 (91%) | 32 (100%) | 72 (97%) | 150 (93%) |  |  |
|  High | 28 (7.2%) | 7 (14%) | 5 (14%) | 3 (8.6%) | 0 (0%) | 2 (2.7%) | 11 (6.8%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.30 (1.10-1.53) | 0.002 |
|  Wild-type | 1544 (95%) | 174 (95%) | 157 (92%) | 124 (91%) | 131 (96%) | 356 (97%) | 602 (96%) |  |  |
|  Mutant | 73 (4.5%) | 9 (5%) | 13 (8%) | 13 (9.5%) | 5 (3.7%) | 10 (2.7%) | 23 (3.7%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.08 (0.96-1.21) | 0.22 |
|  Wild-type | 568 (65%) | 55 (52%) | 44 (62%) | 40 (63%) | 48 (65%) | 144 (67%) | 237 (70%) |  |  |
|  Mutant | 303 (35%) | 50 (48%) | 27 (38%) | 24 (38%) | 26 (35%) | 72 (33%) | 50 (48%) |  |  |
| **Age ≥50 (later-onset)** |  |  |  |  |  |  |  |  |  |
|  **Diagnosed up to 2000** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.18 (2.05-2.33) | <0.001 |
|  Non-MSI-high | 5345 (84%) | 818 (73%) | 559 (64%) | 461 (68%) | 415 (86%) | 1588 (95%) | 1504 (97%) |  |  |
|  MSI-high | 1021 (16%) | 303 (27%) | 313 (36%) | 216 (32%) | 68 (14%) | 76 (4.6%) | 45 (2.9%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 2.20 (2.05-2.36) | <0.001 |
|  Low/negative | 4487 (82%) | 698 (71%) | 450 (58%) | 403 (67%) | 363 (91%) | 1308 (95%) | 1265 (96%) |  |  |
|  High | 970 (18%) | 288 (29%) | 324 (42%) | 199 (33%) | 36 (9.0%) | 74 (5.4%) | 49 (3.7%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 1.84 (1.73-1.95) | <0.001 |
|  Wild-type | 5199 (86%) | 817 (78%) | 580 (69%) | 475 (74%) | 406 (90%) | 1470 (94%) | 1451 (95%) |  |  |
|  Mutant | 868 (14%) | 235 (22%) | 258 (31%) | 168 (26%) | 44 (10%) | 88 (5.6%) | 75 (4.9%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 1.00 (0.95-1.04) | 0.83 |
|  Wild-type | 3663 (67%) | 550 (57%) | 530 (69%) | 421 (72%) | 267 (65%) | 1019 (71%) | 876 (68%) |  |  |
|  Mutant | 1776 (33%) | 421 (43%) | 237 (31%) | 161 (28%) | 144 (35%) | 408 (29%) | 405 (32%) |  |  |
|  **Diagnosed after 2000** |  |  |  |  |  |  |  |  |  |
|  MSI status |  |  |  |  |  |  |  | 2.10 (1.94-2.28) | <0.001 |
|  Non-MSI-high | 3267 (85%) | 400 (74%) | 429 (65%) | 250 (72%) | 227 (90%) | 785 (96%) | 1176 (97%) |  |  |
|  MSI-high | 571 (15%) | 144 (26%) | 236 (35%) | 96 (28%) | 24 (10%) | 36 (4.4%) | 35 (2.9%) |  |  |
|  CIMP status |  |  |  |  |  |  |  | 1.71 (1.58-1.84) | <0.001 |
|  Low/negative | 2381 (82%) | 272 (71%) | 313 (62%) | 181 (69%) | 151 (85%) | 574 (93%) | 890 (93%) |  |  |
|  High | 520 (18%) | 113 (29%) | 192 (38%) | 80 (31%) | 27 (15%) | 44 (7.1%) | 64 (6.7%) |  |  |
|  *BRAF*  |  |  |  |  |  |  |  | 2.15 (1.94-2.38) | <0.001 |
|  Wild-type | 3229 (89%) | 422 (81%) | 456 (72%) | 259 (80%) | 209 (92%) | 741 (96%) | 1142 (99%) |  |  |
|  Mutant | 395 (11%) | 96 (19%) | 173 (28%) | 66 (20%) | 17 (7.5%) | 29 (3.8%) | 14 (1.2%) |  |  |
|  *KRAS*  |  |  |  |  |  |  |  | 0.99 (0.93-1.04) | 0.60 |
|  Wild-type | 2012 (65%) | 217 (54%) | 346 (65%) | 195 (71%) | 136 (69%) | 448 (68%) | 670 (66%) |  |  |
|  Mutant | 1061 (35%) | 187 (46%) | 185 (35%) | 78 (29%) | 60 (31%) | 207 (32%) | 344 (34%) |  |  |

\* Percentage 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).

† Multivariable odds ratio (OR) (with 95% confidence interval) for molecular marker positivity (i.e., MSI-high, CIMP-high, *BRAF* mutation, or *KRAS* mutation) and Ptrend 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) and study (i.e., cohort).

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