Supplementary Methods

Molecular Analyses

DNA was extracted from formalin-fixed paraffin-embedded tissue specimens and mutation status of *KRAS* (codons 12 and 13) (1, 2), *BRAF* (codon 600) (3), and *PIK3CA* (exons 9 and 20) (4) was assessed by Pyrosequencing. Microsatellite instability (MSI) analysis was performed by polymerase chain reaction (PCR) using a panel of 10 microsatellite markers, as previously described (3). MSI-high was defined as instability in ≥30% of the markers, and MSI-low as instability in <30% of the markers (3). We have previously demonstrated that microsatellite stability MSS (no unstable markers) and MSI-low tumors share similar features (3); we therefore included MSI-low cases within the MSS subgroup.

Bisulfite conversion of genomic DNA and real-time PCR (MethyLight) were performed using validated procedures (5). DNA methylation was quantified at eight CpG island methylator phenotype (CIMP)-specific promoters (*CACNA1G, CDKN2A* (*p16*), *CRABP1, IGF2, MLH1, NEUROG1, RUNX3* and *SOCS1*) (6). CIMP-high (methylation at ≥6/8 promoters), CIMP-low (methylation at 1-5/8 promoters), and CIMP-0 (CIMP-negative; no methylated promoters) were defined in accordance with established criteria (6). Long interspersed nucleotide element-1 (LINE-1) methylation was assessed by bisulfite Pyrosequencing (7, 8).

Statistical Methods

For associations between clinical, pathologic, and molecular features, a chi-square test was performed for categorical data. Where table cell counts were <5, and a chi-square test may

have been unreliable, we computed a *P* value using a randomization test of independence. An ANOVA was used to compare means between subgroups for continuous variables. To account for multiple hypothesis testing in associations between clinical, pathologic, and molecular features, the *P* value for significance was adjusted by Bonferroni correction to P=0.0042 (=0.05/12). When testing our main hypothesis, on prognostic associations of MSI/*BRAF* subgroups, we did not adjust for multiple testing; results were, nonetheless, interpreted cautiously. The Kaplan-Meier method and log-rank test were used to estimate survival distribution for molecular subgroups. Cases were observed until death, or January 1st 2011, whichever came first. Median follow-up time was calculated, including censored and uncensored cases. For analyses of colorectal cancer-specific mortality, deaths as a result of other causes were censored. We demonstrated that censoring other causes of death did not substantially bias our cause-specific mortality analyses by confirming that cumulative incidence functions for colorectal-cancer-related failure yielded similar results to the Kaplan-Meier estimator (P<0.001 by Gray's test of equality of cumulative incidence) (**Supplementary Figure 3**).

Cox proportional hazards regression models were used to compute mortality hazard ratios (HR) for molecular subgroups. Stratification by disease stage (I, II, III, IV or unknown) (2) was performed using the "strata" option in the SAS "proc phreg" command. Multivariable stage-stratified models were used to adjust for confounding by clinical, pathological, and other tumor molecular features. Variables included in the multivariable models were selected separately for colorectal cancer-specific survival and overall survival using a backward elimination procedure with a threshold of P=0.10. Variables initially included in the selection procedure were: sex, age at diagnosis (continuous), year of diagnosis (continuous), body mass index (BMI; ≥30 vs. <30

kg/m^2), family history of colorectal cancer in any first-degree relative (present vs. absent), tumor location (proximal vs. distal), tumor differentiation (well-moderate vs. poor), MSI (high vs. low/MSS), CIMP (high vs. low/0), LINE-1 methylation (continuous), and *KRAS*, *BRAF*, and *PIK3CA* mutation (present vs. absent). Cases with missing information in any of the categorical covariates [BMI (0.2% missing); tumor location (0.6%); tumor differentiation (0.6%); CIMP (6.4%); *KRAS* (0.3%); *PIK3CA* (7.5%)] were assigned to the majority category in order to maximise statistical power and improve the efficiency of the multivariable models. We performed a sensitivity analysis (N=1244) by excluding cases with missing covariate information from the multivariable Cox regression models, and confirmed that this did not substantially alter the results obtained (data not shown). For family history and weight, the most recently updated information from the questionnaire cycle preceding diagnosis was used. In addition to MSI status and *BRAF* mutation, covariates selected for inclusion in the colorectal cancer-specific model were age, year of diagnosis, BMI, tumor differentiation, and LINE-1 methylation. The same covariates, with the exception of LINE-1 methylation, were selected for inclusion in the overall survival model. Interaction was assessed by the Wald test on the interaction term that was the cross-product of the variables of interest (*BRAF* and MSI). We observed evidence of non-proportionality in the hazards for MSI/*BRAF* subgroups over time using time-varying covariates; however, using the Schoenfeld test, the hazards were proportional for most of the follow-up period, from around 2 years onward, supporting use of Cox regression models. In addition, our Cox regression data were essentially in agreement with Kaplan-Meier analysis data.

REFERENCES

1. Ogino S, Kawasaki T, Brahmandam M*, et al.* Sensitive sequencing method for KRAS mutation detection by Pyrosequencing. *J Mol Diagn.* 2005;7(3):413-421.

2. Edge SB, Byrd DR, Compton CC*, et al. AJCC cancer staging manual*. 7th ed. New York (NY): Springer; 2010.

3. Ogino S, Kawasaki T, Kirkner GJ*, et al.* CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. *J Mol Diagn.* 2006;8:582-588.

4. Liao X, Morikawa T, Lochhead P*, et al.* Prognostic Role of PIK3CA Mutation in Colorectal Cancer: Cohort Study and Literature Review. *Clin Cancer Res.* 2012;18(8):2257-2268.

5. 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(2):209-217.

6. Nosho K, Irahara N, Shima K*, et al.* Comprehensive biostatistical analysis of CpG island methylator phenotype in colorectal cancer using a large population-based sample. *PLoS ONE.* 2008;3(11):e3698.

7. Ogino S, Nosho K, Kirkner GJ*, et al.* A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer. *J Natl Cancer Inst.* 2008;100:1734-1738.

8. Irahara N, Nosho K, Baba Y*, et al.* Precision of Pyrosequencing assay to measure LINE-1 methylation in colon cancer, normal colonic mucosa and peripheral blood cells. *J Mol Diagn.* 2010;12(2):177-183.