Supplemental Table 1: Results of Multivariate Models of Association between Clinical Endpoints and Methylation Markers Assayed from Stool, Adjusting for Clinical Variables

Model	P-values by endpoint		
	Cancer[5]	Neoplasia	Dysplasia
mBMP3	0.04	0.004	0.01
Age[1]	0.98	0.87	0.61
Age x *mBMP3*	0.86	0.92	0.91
mBMP3	0.009	0.004	0.03
Sex	0.03	0.11	0.31
Sex x *mBMP3*	0.07	0.09	0.33
mBMP3	0.45	0.006	0.01
IBD Duration[2]	0.05	0.47	0.12
IBD Duration x *mBMP3*	0.07	0.85	0.67
mBMP3	0.02	0.02	0.06
IBD Extent[3]	Unstable	0.03	0.09
IBD Extent x *mBMP3*	Unstable	0.08	0.25
mBMP3	0.03	0.25	0.26
PSC[4]	Unstable	0.36	0.28
PSC x *mBMP3*	Unstable	0.36	0.35
mVIM	0.01	0.01	0.05
Age	0.37	0.96	0.97
Age x *mVIM*	0.50	0.94	0.93
mVIM	0.02	0.002	0.02
Sex	0.03	0.03	0.33
Sex x *mVIM*	Unstable	0.02	0.08
mVIM	0.67	0.02	0.04
IBD Duration	0.12	0.20	0.85
IBD Duration x *mVIM*	0.34	0.87	0.30
mVIM	0.03	0.002	0.06
IBD Extent	Unstable	0.001	0.08
IBD Extent x *mVIM*	Unstable	0.01	0.08

mVIM	0.04	0.33	0.35
PSC	0.98	0.44	0.38
PSC x *mVIM*	Unstable	0.44	0.41
mEYA4	0.02	0.01	0.02
Age	0.84	0.37	0.95
Age x *mEYA4*	0.93	0.24	0.20
mEYA4	0.01	0.003	0.03
Sex	0.38	0.48	0.48
Sex x *mEYA4*	Unstable	0.57	0.57
mEYA4	0.19	0.008	0.03
IBD Duration	0.12	0.12	0.13
IBD Duration x *mEYA4*	0.49	0.80	0.59
mEYA4	0.02	0.01	0.14
IBD Extent	0.38	0.04	0.11
IBD Extent x *mEYA4*	Unstable	0.55	0.90
mEYA4	0.01	0.07	0.12
PSC	Unstable	0.92	0.72
PSC x *mEYA4*	Unstable	0.98	0.73
mNDRG4	0.03	0.008	0.02
Age	0.82	0.47	0.96
Age x *mNDRG4*	0.17	0.07	0.14
mNDRG4	0.003	0.003	0.02
Sex	0.11	0.52	0.41
Sex x *mNDRG4*	Unstable	0.39	0.62
mNDRG4	0.67	0.01	0.03
IBD Duration	0.08	0.03	0.07
IBD Duration x *mNDRG4*	0.18	0.56	0.95
mNDRG4	0.01	0.01	0.30
IBD Extent	0.36	0.03	0.13
IBD Extent x *mNDRG4*	Unstable	0.52	0.69
mNDRG4	0.01	0.06	0.12
PSC	0.98	0.83	0.57
PSC x *mNDRG4*	Unstable	0.77	0.52

1. Age in years at time of study consent
2. Years since inflammatory bowel disease (IBD) diagnosis

3. Left-sided colitis versus colitis proximal to splenic flexure
4. Presence or absence of comorbid primary sclerosing cholangitis (PSC)
5. Multivariate regression models with unstable terms were repeated, excluding the unstable variable(s)

Supplemental Methods

Tissue Study

DNA Extraction: Using a modified Gentra (Gentra Systems Inc., Minneapolis, MN) protocol, DNA extracted from paraffin-embedded tissues was suspended in TE (10 mM Tris/ 0.1 mM EDTA, Integrated DNA Technologies, Coralville, IA). Quantification of total DNA was performed using the Picogreen assay (Invitrogen, Portland, OR).[34]

Mutation Marker Gene Sequencing: Candidate exons on *APC, p53, K-ras, BRAF* and *PIK3CA* were amplified in a real-time iCycler (BioRad, Hercules, CA) using real-time PCR reactions, performed with sense and antisense primers, IQ Supermix polymerase kit (BioRad) and 10 ng of genomic DNA. Products were run on a 2% agarose gel to confirm the presence of a single band and then cleaned with ExoSAP-IT (Affymetrix, Santa Clara, CA). The 14 exons of interest were bidirectionally sequenced on all 50 specimens on an ABI PRISM 3730xl DNA analyzer (Applied Biosystems Inc, Foster City, CA). Sequences were screened for mutations using Mutation Surveyor (SoftGenetics, State College, PA) software and then compared to the National Center for Biotechnology Information database of single-nucleotide polymorphisms (dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/) to exclude common variants.

Real-Time Methylation-Specific PCR (MSP): DNA was bisulfite treated using the EZ DNA Methylation Kit (Zymo Research, Orange, CA). Primers were designed to target the bisulfite-modified methylated sequences of gene promoters (IDT, Coralville, IA). The *ß-actin* gene was quantified with real-time PCR using primers and probe recognizing bisulfite-converted sequence as a reference.

Stool Study

Stool Collection: Using a plastic bucket device mounted on the toilet seat, whole stools were collected and then stabilized with buffer solution and sealed with a water-tight lid. Upon laboratory receipt, stools were homogenized, aliquoted, and frozen at -80C until assayed.[37, 38]

Sequence-specific gene capture: Stool samples were weighed and diluted 1:5 with additional buffer before incubation with polyvinylpyrrolidone (Crosby & Baker, Westport, MA) to remove PCR inhibitors. A 2-gram equivalent of stool supernatant was used for multiplex capture of 4 gene targets (*β-actin, VIM, EYA4, BMP3* and *NDRG4*). Sodium chloride and guanidine thiocyanate (Sigma, St. Louis, MO) denaturation buffer were added to clarified stool supernatant and heated in a water bath before incubation and room temperature hybridization with carboxylic acid–coated capture beads with amino conjugated oligonucleotides complementary to target sequences (IDT). A 3-step wash in MOPS buffer was performed prior to heated tRNA buffer elution.

Assay of Methylated Markers: Quantitative allele-specific real-time target and signal amplification (QuARTS) reactions were performed on Roche 480 LightCyclers (Indianapolis, IN) using sets of primers, detection probes and invasive oligonucleotides (FAM, Hologic,Madison WI), fluorescence resonance energy transfers (FRETs), Cleavase 2.0 (Hologic), GoTaq DNA polymerase (Promega, Madison, WI), 10 mM MOPS, 7.5 mM MgCl2, and 250 µM of each dNTP for *β-actin, mBMP3, mVIM* and *mNDRG4* genes. Bisulfite-treated CpGenome™ Universal methylated DNA (Millipore) and human genomic DNA (Novogen, Oakville, Canada) were used as positive and

negative controls. Each plate contained standards made of engineered plasmids, positive and negative controls, and water blanks. Standard curves were made of 10-fold serially diluted engineered plasmids with corresponding gene inserts to calculate the copy number of each marker based on an amplification efficiency of 1.95.

All oligonucleotide sequences for the tissue and the stool studies are available on request.

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls
		(*b) Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
Outcome data	15*	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.