Supplementary Table 4. Most likely target gene(s) at the 13 new loci identified across the CRC case subgroup analyses.

Locus	Putative target gene(s)	Expression-based linking (GTEx V8)	Biological relevance, experimental functional evidence, somatic alterations, familial syndromes
1p31.1	*PTGER3*	*PTGER3* (7 tissues)	*PTGER3* encodes Prostaglandin E Receptor 3, a receptor for prostaglandin E2 (PGE2), a potent pro-inflammatory metabolite that is biosynthesized by Cyclooxygenase-2 (COX-2). COX-2 plays a critical role in mediating inflammatory responses that lead to epithelial malignancies and its expression is induced by NF-κβ and TNF-α. The anti-inflammatory activity of nonsteroidal anti-inflammatory drugs (NSAIDs) such as aspirin and ibuprofen operates mainly through COX-2 inhibition, and long-term NSAID use decreases incidence and mortality from CRC.[1] Prostaglandin E2 (PGE2) is required for the activation of β-catenin by Wnt in stem cells,[2] and promotes colon cancer cell growth.[3] Prostaglandin E Receptor 3 plays an important role in suppression of cell growth and its downregulation was shown to enhance colon carcinogenesis.[4] Hypermethylation may contribute to its downregulation in colon cancer.[4]
2q21.3	*LCT*		Lead SNP rs1446585 is in strong LD with the functional SNP rs4988235 (LD $r^2 = 0.854$) in the *cis*-regulatory element of the lactase gene. In Europeans, the rs4988235 genotype determines the autosomal dominant lactase persistence phenotype, or the ability to digest the milk sugar lactose in adulthood. The allele determining lactase persistence (T) is associated with a decreased risk of CRC. This is consistent with a previous candidate study that reported a significant association between low lactase activity defined by the CC genotype and CRC risk in the Finnish population.[5] The protective effect conferred

			by the lactase persistence genotype is likely mediated by dairy products and calcium which are known protective factors for CRC.[6] Consistent with a dominant model, associations for rs1446585 and rs4988235 became more significant when tested assuming a dominant model with *P*-values of 4.4×10^{-11} and 1.4×10^{-9}, respectively (see main text).
3p22.2	*MLH1*	*MLH1* (14 tissues)	Previous candidate gene studies have reported strong and robust associations between the common, *MLH1* gene promoter region and lead SNP rs1800734, and sporadic CRC cases with high microsatellite instability (MSI-H) status with consistent direction of effects.[7,8] Rare deleterious nonsynonymous mutations in the DNA mismatch repair (MMR) gene *MLH1* are a cause of Lynch syndrome (OMIM #609310). The risk allele of the likely causal SNP rs1800734 showed a strong association with *MLH1* promoter hypermethylation and loss of MLH1 protein in CRC tumors.[8] The mechanisms of *MLH1* promoter hypermethylation and subsequent gene silencing may account for most sporadic CRC tumors with defective DNA MMR and MSI-H.[9]
3p21.2	*STAB1; TLR9; NISCH*	*STAB1* (10 tissues); *TLR9* (3 tissues); *NISCH* (4 tissues)	This signal is located in a gene dense region. The Stabilin 1 (*STAB1*) gene encodes an endocytotic scavenger receptor expressed in a number of cell types, including activated macrophages in human malignancies.[10] A rare missense variant in *STAB1* has previously shown to be strongly associated with serum lactate dehydrogenase (LDH) levels,[11] a widely used marker of tissue damage, affirming a link between STAB1 and the clearance of products of cell lysis through the mononuclear phagocytic system. Human Protein Atlas data based on The Cancer Genome Atlas (TCGA) show that *STAB1* expression is an unfavorable prognostic marker for CRC (logrank test *P*=0.0008, based on maximally separated Kaplan-Meier curves; *n* samples=597). Lead SNP rs353548 is

			located in an intron of the toll like receptor 9 (*TLR9*) gene which could also be involved. This key component of innate and adaptive immunity is a drug target for many immune-mediated diseases, and the antagonist drug hydroxychloroquine is included in chemotherapy combination clinical trials for colorectal carcinoma (ClinicalTrials.gov Identifier: NCT01006369). The Nischarin (*NISCH*) gene encodes an α5 integrin-binding protein and may be a tumor suppressor gene that limits breast cancer progression.[12] Nischarin inhibits Rac-induced cell migration and invasion in breast and colon epithelial cells.[13]
5q32	*CDX1*		The intestine-specific transcription factor caudal-type homeobox 1(*CDX1*) encodes a key regulator of differentiation of enterocytes in the normal intestine and of CRC cells. CDX1 is central to the capacity of colon cells to differentiate and promotes differentiation by repressing the polycomb complex protein BMI1 which promotes stemness and self-renewal. Colonic crypt cells express BMI1 but not CDX1. The repression of BMI1 is mediated by microRNA-215 which acts as a target of CDX1 to promote differentiation and inhibit stemness.[14] Consistent with this view, CDX1 has been shown to inhibit human colon cancer cell proliferation by blocking β-catenin/T-cell factor transcriptional activity.[15]
7q32.3	*KLF14; LINC00513*	*LINC00513* (transverse colon + 2 tissues)	The Krüppel-like factor 14 (*KLF14*) gene is a strong candidate involved in TGF-β signaling. We previously reported loci at known CRC oncogene *KLF5* and at *KLF2*.[16] The imprinted gene *KLF14* shows monoallelic maternal expression, and is induced by TGF-β to transcriptionally corepress the TGF-beta receptor II (*TGFBR2*) gene.[17] A cis-eQTL for *KLF14*, that is uncorrelated with our lead SNP rs73161913, acts as a master

			regulator related to multiple metabolic phenotypes,[18,19] and an independent variant in this region has been associated to basal cell carcinoma.[20] The signal overlaps with an eQTL for the lncRNA gene *LINC00513* which may be involved in the regulation of *KLF14* expression.
10q23.31	*PANK1; KIF20B*	*PANK1* (transverse colon + 3 tissues); *KIF20B* (transverse colon + 7 tissues)	At 10q23.31, GTEx data show that the lead SNP rs7071258 is an eQTL in transverse colon tissue for genes Pantothenate Kinase 1 (*PANK1*) and Kinesin Family Member 20B (*KIF20B*). The enzyme encoded by *PANK1* catalyzes the rate-limiting reaction in the biosynthesis of coenzyme A and may play a role in tumor metabolism.[21] *KIF20B* has been suggested to play an oncogenic role in bladder carcinogenesis.[22] *KIF20B* missense variant rs34354493 (canonical transcript, p.Lys1609Glu) is in high LD with the lead variant (r^2=0.90) and is predicted to be deleterious by multiple algorithms (CADD, DANN, Polyphen, SIFT).
14q22.1	*PYGL; NIN; ABHD12B*	*PYGL* (transverse colon + 12 tissues); *ABHD12B* (transverse colon + 8 tissues); *NIN* (transverse	GTEx data show that, in gastrointestinal tissues, the lead SNP is a *cis*-eQTL co-regulating expression of genes *PYGL*, *ABHD12B*, and *NIN*. Glycogen Phosphorylase L (*PYGL*) is the strongest candidate. We recently identified and replicated an association between genetically predicted *PYGL* expression and CRC risk in a transcriptome-wide association study that used transverse colon tissue transcriptomes and genotypes from GTEx to construct prediction models.[23] Favaro *et al.* showed that this glycogen metabolism gene plays an important role in sustaining proliferation and preventing premature senescence in hypoxic cancer cells.[24] In different cancer cells lines, silencing of *PYGL*, expression of which is induced by exposure to hypoxia, led to increased glycogen accumulation and

		colon + 2 tissues)	increased reactive oxygen species levels that contributed to p53-dependent induction of senescence and impaired tumorigenesis.[24]
14q32.12	*RIN3*	*RIN3* (transverse colon + 11 tissues)	Lead SNP rs61975764 is an eQTL for gene Ras And Rab Interactor 3 (*RIN3*) in colon tissue, the risk allele G being associated with decreased expression. RIN3 functions as a RAB5 and RAB31 guanine nucleotide exchange factor involved in endocytosis.[25,26]
14q32.2	*BCL11B*		The lead SNP rs80158569 of this highly localized proximal colon-specific association signal is located in a normal colonic crypt enhancer region and overlaps with multiple transcription factor binding sites, making it a strong functional candidate. The nearby gene *BCL11B* encodes a transcription factor that is required for normal T cell development,[27,28] and that has been identified as a SWI/SNF complex subunit.[29] *BCL11B* acts as a haploinsufficient tumor suppressor in T-cell acute lymphoblastic leukemia (T-ALL).[30,31] Experimental work reported by Sakamaki *et al.* suggests that impairment of Bcl11b promotes intestinal tumorigenesis in mice and humans through deregulation of the β-catenin pathway.[32]
19p13.3	*STK11; SBNO2*		This signal is located in a gene-dense region. Lead SNP rs62131228 is intronic to gene Strawberry notch homologue 2 (*SBNO2*), a transcriptional corepressor of NF-κβ in macrophages that plays a role in the STAT3-regulated anti-inflammatory signaling pathway.[33] The nearby tumor suppressor gene Serine/Threonine Kinase 11 (*STK11*) is an especially plausible candidate effector gene. Mutations in this gene cause Peutz-Jeghers syndrome (OMIM #175200), an autosomal dominant disorder characterized by the growth

			of hamartomatous gastrointestinal polyps and an increased risk of various neoplasms.[34,35]
20q13.31	*BMP7*	*BMP7* (3 tissues)	The Bone Morphogenetic Protein 7 (BMP7) gene is a strong candidate. In normal intestinal cell crypts, various gradients of TGF-β family members interact with the antagonistic Wnt signaling pathway to maintain homeostasis. Members of the TGF-β family, including several bone morphogenetic proteins (BMPs), frequently have somatic mutations in sporadic CRC tumors, have been implicated by GWASs, and germline mutations are causative for familial CRC syndromes.[36] BMP7 signaling in *TGFBR2*-deficient stromal cells promotes epithelial carcinogenesis through SMAD4-mediated signaling.[37] In CRC tumors, *BMP7* expression correlates with parameters of pathological aggressiveness such as liver metastasis and poor prognosis.[38]
22q13.31	*FAM118A; FBLN1*	*FAM118A* (transverse colon + 40 tissues)	GTEx data show that the lead SNP rs736037 is an eQTL for gene *FAM118A* in many tissues, including transverse colon. The function of *FAM118A* is poorly understood. *FAM118A* missense variant rs6007594 (canonical transcript, p.Arg239His) is in high LD with the lead variant rs736037 (r^2=0.96) and is predicted to be deleterious by multiple algorithms (CADD, DANN, Polyphen). The protein encoded by the nearby Fibulin 1 (*FBLN1*) gene plays a role in the organization and function of the extracellular matrix and basement membranes. *FBLN1* has been implicated in tumor-related processes and both oncogenic and tumor-suppressive properties have been described for this protein.[39] Other genes in the region are no obvious candidates.

References

1 Jänne PA, Mayer RJ. Chemoprevention of colorectal cancer. *N Engl J Med* 2000;**342**:1960–8. doi:10.1056/NEJM200006293422606
2 Goessling W, North TE, Loewer S, *et al.* Genetic interaction of PGE2 and Wnt signaling regulates developmental specification of stem cells and regeneration. *Cell* 2009;**136**:1136–47. doi:10.1016/j.cell.2009.01.015
3 Castellone MD, Teramoto H, Williams BO, *et al.* Prostaglandin E2 promotes colon cancer cell growth through a Gs-axin-beta-catenin signaling axis. *Science* 2005;**310**:1504–10. doi:10.1126/science.1116221
4 Shoji Y, Takahashi M, Kitamura T, *et al.* Downregulation of prostaglandin E receptor subtype EP3 during colon cancer development. *Gut* 2004;**53**:1151–8. doi:10.1136/gut.2003.028787
5 Rasinperä H, Forsblom C, Enattah NS, *et al.* The C/C-13910 genotype of adult-type hypolactasia is associated with an increased risk of colorectal cancer in the Finnish population. *Gut* 2005;**54**:643–7. doi:10.1136/gut.2004.055939
6 World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Diet, nutrition, physical activity and colorectal cancer. *Available at dietandcancerreport.org*
7 Raptis S, Mrkonjic M, Green RC, *et al.* MLH1 -93G>A promoter polymorphism and the risk of microsatellite-unstable colorectal cancer. *J Natl Cancer Inst* 2007;**99**:463–74. doi:10.1093/jnci/djk095
8 Mrkonjic M, Roslin NM, Greenwood CM, *et al.* Specific variants in the MLH1 gene region may drive DNA methylation, loss of protein expression, and MSI-H colorectal cancer. *PLoS One* 2010;**5**:e13314. doi:10.1371/journal.pone.0013314
9 Cunningham JM, Christensen ER, Tester DJ, *et al.* Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res* 1998;**58**:3455–60.
10 Schönhaar K, Schledzewski K, Michel J, *et al.* Expression of stabilin-1 in M2 macrophages in human granulomatous disease and melanocytic lesions. *Int J Clin Exp Pathol* 2014;**7**:1625–34.
11 Kristjansson RP, Oddsson A, Helgason H, *et al.* Common and rare variants associating with serum levels of creatine kinase and lactate dehydrogenase. *Nat Commun* 2016;**7**:10572. doi:10.1038/ncomms10572
12 Baranwal S, Wang Y, Rathinam R, *et al.* Molecular characterization of the tumor-suppressive function of nischarin in breast cancer. *J Natl Cancer Inst* 2011;**103**:1513–28. doi:10.1093/jnci/djr350
13 Alahari SK. Nischarin inhibits Rac induced migration and invasion of epithelial cells by affecting signaling cascades involving PAK. *Exp Cell Res* 2003;**288**:415–24.
14 Jones MF, Hara T, Francis P, *et al.* The CDX1-microRNA-215 axis regulates colorectal cancer stem cell differentiation. *Proc Natl Acad Sci USA* 2015;**112**:E1550-8. doi:10.1073/pnas.1503370112
15 Guo R-J, Huang E, Ezaki T, *et al.* Cdx1 inhibits human colon cancer cell proliferation by reducing beta-catenin/T-cell factor transcriptional activity. *J Biol Chem* 2004;**279**:36865–75. doi:10.1074/jbc.M405213200
16 Huyghe JR, Bien SA, Harrison TA, *et al.* Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet*

2019;**51**:76–87. doi:10.1038/s41588-018-0286-6
17 Truty MJ, Lomberk G, Fernandez-Zapico ME, *et al.* Silencing of the transforming growth factor-beta (TGFbeta) receptor II by Kruppel-like factor 14 underscores the importance of a negative feedback mechanism in TGFbeta signaling. *J Biol Chem* 2009;**284**:6291–300. doi:10.1074/jbc.M807791200
18 Small KS, Hedman AK, Grundberg E, *et al.* Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nat Genet* 2011;**43**:561–4. doi:10.1038/ng.833
19 Small KS, Todorčević M, Civelek M, *et al.* Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nat Genet* 2018;**50**:572–80. doi:10.1038/s41588-018-0088-x
20 Stacey SN, Sulem P, Masson G, *et al.* New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009;**41**:909–14. doi:10.1038/ng.412
21 Wang S-J, Yu G, Jiang L, *et al.* p53-Dependent regulation of metabolic function through transcriptional activation of pantothenate kinase-1 gene. *Cell Cycle* 2013;**12**:753–61. doi:10.4161/cc.23597
22 Kanehira M, Katagiri T, Shimo A, *et al.* Oncogenic role of MPHOSPH1, a cancer-testis antigen specific to human bladder cancer. *Cancer Res* 2007;**67**:3276–85. doi:10.1158/0008-5472.CAN-06-3748
23 Bien SA, Su Y-R, Conti DV, *et al.* Genetic variant predictors of gene expression provide new insight into risk of colorectal cancer. *Hum Genet* 2019;**138**:307–26. doi:10.1007/s00439-019-01989-8
24 Favaro E, Bensaad K, Chong MG, *et al.* Glucose utilization via glycogen phosphorylase sustains proliferation and prevents premature senescence in cancer cells. *Cell Metab* 2012;**16**:751–64. doi:10.1016/j.cmet.2012.10.017
25 Kajiho H, Saito K, Tsujita K, *et al.* RIN3: a novel Rab5 GEF interacting with amphiphysin II involved in the early endocytic pathway. *J Cell Sci* 2003;**116**:4159–68. doi:10.1242/jcs.00718
26 Kajiho H, Sakurai K, Minoda T, *et al.* Characterization of RIN3 as a guanine nucleotide exchange factor for the Rab5 subfamily GTPase Rab31. *J Biol Chem* 2011;**286**:24364–73. doi:10.1074/jbc.M110.172445
27 Avram D, Califano D. The multifaceted roles of Bcl11b in thymic and peripheral T cells: impact on immune diseases. *J Immunol* 2014;**193**:2059–65. doi:10.4049/jimmunol.1400930
28 Punwani D, Zhang Y, Yu J, *et al.* Multisystem Anomalies in Severe Combined Immunodeficiency with Mutant BCL11B. *N Engl J Med* 2016;**375**:2165–76. doi:10.1056/NEJMoa1509164
29 Kadoch C, Hargreaves DC, Hodges C, *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* 2013;**45**:592–601. doi:10.1038/ng.2628
30 Gutierrez A, Kentsis A, Sanda T, *et al.* The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. *Blood* 2011;**118**:4169–73. doi:10.1182/blood-2010-11-318873
31 Neumann M, Vosberg S, Schlee C, *et al.* Mutational spectrum of adult T-ALL. *Oncotarget* 2015;**6**:2754–66. doi:10.18632/oncotarget.2218
32 Sakamaki A, Katsuragi Y, Otsuka K, *et al.* Bcl11b SWI/SNF-complex subunit modulates intestinal adenoma and regeneration

after γ-irradiation through Wnt/β-catenin pathway. *Carcinogenesis* 2015;**36**:622–31. doi:10.1093/carcin/bgv044

33 El Kasmi KC, Smith AM, Williams L, *et al.* Cutting edge: A transcriptional repressor and corepressor induced by the STAT3-regulated anti-inflammatory signaling pathway. *J Immunol* 2007;**179**:7215–9. doi:10.4049/jimmunol.179.11.7215

34 Hemminki A, Tomlinson I, Markie D, *et al.* Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nat Genet* 1997;**15**:87–90. doi:10.1038/ng0197-87

35 Hemminki A, Markie D, Tomlinson I, *et al.* A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature* 1998;**391**:184–7. doi:10.1038/34432

36 Jung B, Staudacher JJ, Beauchamp D. Transforming Growth Factor β Superfamily Signaling in Development of Colorectal Cancer. *Gastroenterology* 2017;**152**:36–52. doi:10.1053/j.gastro.2016.10.015

37 Eikesdal HP, Becker LM, Teng Y, *et al.* BMP7 Signaling in TGFBR2-Deficient Stromal Cells Provokes Epithelial Carcinogenesis. *Mol Cancer Res* 2018;**16**:1568–78. doi:10.1158/1541-7786.MCR-18-0120

38 Motoyama K, Tanaka F, Kosaka Y, *et al.* Clinical significance of BMP7 in human colorectal cancer. *Ann Surg Oncol* 2008;**15**:1530–7. doi:10.1245/s10434-007-9746-4

39 Fontanil T, Mohamedi Y, Cobo T, *et al.* Novel associations within the tumor microenvironment: fibulins meet adamtss. *Front Oncol* 2019;**9**:796. doi:10.3389/fonc.2019.00796