Single Nucleotide Polymorphisms in Colorectal Cancer: Associations with Tumor Site and TNM Stage

Ioan Nicolae Mates[1]*, Viorel Jinga[2]*, Irma Eva Csiki[3], Dana Mates[3], Daniela Dinu[1], Adrian Constantin[1], Mariana Jinga[4]

1) St Mary General and Esophageal Surgery Clinic, Carol Davila University of Medicine and Pharmacy; 2) Th Burghele Urology Clinic, Carol Davila University of Medicine and Pharmacy; 3) National Institute of Public Health, Bucharest; 4) Internal Medicine and Gastroenterology Clinic, Carol Davila University of Medicine and Pharmacy, Bucharest, Romania

Abstract

Background & aim. Colon tumor carcinogenesis and rectal tumor carcinogenesis have each been associated with different genetic features, but data are still controversial and are insufficient to support their distinct molecular biology. Recently, genome-wide association studies (GWAS) have also found heterogeneity in colorectal cancer (CRC) risks based on population ethnicity and tumor features. Several single nucleotide polymorphism (SNP) markers are described in the literature as having site and/or stage specificity, including rs10795668, rs3802842, rs6983267, and rs4939827. Replication of initial findings in different ethnic groups by independent studies is required to unravel the population-specific differences in risk. **Methods.** We examined whether inherited risk variants at rs10795668, rs3802842, rs6983267, and rs4939827 exerted a differential effect on colon and rectal cancers in a Romanian hospital based series of 153 CRC cases and 182 non-affected control subjects prospectively recruited between 2007 and 2010. **Results.** Rectal tumors were significantly associated with rs4939827 (OR = 4.85, P = 0.002) and rs6983267 (OR = 3.00, P = 0.036), suggesting that carriers of risk alleles at these loci had increased susceptibility to development of rectal cancer rather than colon cancer. Carrying the C allele at rs3802842 appeared to be associated with a lower risk for rectal tumors in our dataset. We found no association between genotypes and tumor aggressiveness as reflected by TNM staging. **Conclusions.** The associations between SNPs, and tumor site and staging remain to be further clarified. Our results should be considered cautiously, but may be taken into acount in future, larger epidemiological studies.

*I N Mates and V Jinga equally contributed to this study

Received: 18.10.2011 Accepted: 20.01.2012
J Gastrointestin Liver Dis
March 2012 Vol. 21 No 1, 45-52
Address for correspondence: Ioan Nicolae Mates, PhD, Lecturer
"St. Mary" Clinical Hospital, General and Esophageal Surgery Clinic
Bucharest, Romania
Email: ioan.mates@gmail.com

Key words

Colorectal cancer risk – single nucleotide polymorphisms – tumor site – TNM staging.

Introduction

Colorectal cancer (CRC) is one of the most common neoplasia and the third leading cause of cancer deaths, preceded only by lung cancer in male and breast cancer in female patients [1]. High penetrance mutations, such of those of the APC and MMR genes, account for less than 5% of cases in the pathogenesis of CRC. Another 8.3% cases arise in families with two affected first or second degree relatives; in this group, mildly or moderately penetrant alleles could explain the familial aggregation [2, 3]. It is expected that the remaining proportion of inherited susceptibility is likely to be explained by low-risk variants.

Genome-wide association studies (GWAS) made possible the genotyping of hundreds of thousands of single nucleotide polymorphisms (SNPs) that tag linkage disequilibrium (LD) blocks in the genome. So far, the tagging SNPs with the strongest reported association signals for CRC are rs6983267 at 8q24 [4], rs4779584 at 15q13 [5], rs4939827 at 18q21 [6], rs16892766 at 8q23 [7], rs10795668 at 10p14 [7], rs3802842 at 11q23 [8], rs4444235 at 14q22 [9], rs9929218 at 16q22 [9], rs10411210 at 19q13 [9], and rs961253 at 20p12 [9].

Despite advances in discovering low-risk variants involved in CRC etiology, considerable challenges lie ahead, such as identifying causative variants and elucidating their functional role, understanding and determining the contribution of these loci to the cancer risk in different ethnic groups, and demonstrating their role in morphogenetic pathways of colon and rectal cancers. In the genetic architecture of CRC, the number and distribution of low-risk alleles in any given population depend on several parameters (mutation rate, genetic selection, population demography), so it is possible that variants confer risk in one specific population but not in another [10, 11]. Therefore, it may require replication of initial findings in different ethnic groups by independent studies to unravel the population-specific differences in risk.

Recent reports suggest differential carcinogenesis of sporadic cancers of right colon, left colon (including rectosigmoid junction), and rectum. Other data suggest different prognosis and outcome leading to a need for individualized therapy [12-14]. Data concerning site-specific genetic pathways exist, but are still controversial and more evidence is required to support differences between the molecular biology of colon and rectal cancers.

There are two major tumorigenic pathways leading to CRC. The first is driven by chromosomal instability (CIN), leading to a progression from normal mucosa to adenoma and carcinoma. Mutations accumulate in the KRAS oncogene and tumor-suppressor genes, including APC on 5q, TP53 on 17p, and SMAD4 on 18q [15]. The few studies that have analyzed rectal cancers indicate that 80-98% arise through this CIN pathway [16, 17]. The second pathway is characterized by mutations in mismatch-repair genes (MMR). If somatic cells are affected, microsatellite instability (MSI) is responsible for sporadic tumors [18]. MSI is reported in <10% of rectal cancers, and it is still unknown which target genes are involved; notably, most of the MSI observed in MMR-deficient cells are bystander events that play no role in carcinogenesis [19].

By evaluating the independent effects of the common variants (SNPs) on CRC risk, GWAS also found heterogeneity in associations based on tumor location and tumor stage, providing additional evidence for different carcinogenic pathways across CRC subgroups. Several studies have published genotype-phenotype analyses for the loci 8q24, 18q21, 10p14, and 11q23, and sex, tumor site, tumor stage, age at onset, and family history [20-22]. Some reported significant associations with tumor site, stage, and morphology [7, 8, 20-23].

To investigate the hypothesis of different patterns of CRC susceptibility across populations and phenotypic subgroups, we aimed to assess whether inherited risk variants at rs10795668, rs3802842, rs6983267, and rs4939827 exerted a differential effect on colon and rectal cancers, and to compare our results with those reported in other populations. To our knowledge, this is the first study of the genotype-phenotype association for these loci in a Romanian population sample.

Material and methods

Study population

Our hospital-based investigation comprised a series of 153 CRC cases and 182 non-affected control subjects. All subjects were of self-reported Romanian ethnicity and were residents in the South-Eastern part of Romania. Cases included unrelated men and women, each with a histologically confirmed diagnosis of CRC. They were each recruited between 2007 and 2010, during hospitalization for CRC surgical treatment at the St. Mary Clinic of General and Esophageal Surgery, Carol Davila University of Medicine of Pharmacy, Bucharest, Romania. Controls consisted of patients from the same clinic, as well as from Th. Burghele Clinic, Carol Davila University of Medicine of Pharmacy, Bucharest, Romania, who were hospitalized for various conditions but with no evidence of cancer. Control subjects were randomly selected based on admission records, and frequency-matched to cases based on age groups and sex. All eligible subjects were invited to participate in the study and to donate a blood sample; a 9-ml whole-blood sample was collected from each subject using EDTA vacutainers. Blood tubes were barcode-labeled and stored at −80°C before shipment on dry ice to deCode Genetics (Reykjavik, Iceland) for genotyping.

Demographic and lifestyle information, and medical and family cancer history data were collected by trained interviewers via direct interview using standardized questionnaires. Clinical data were abstracted from hospital medical records and histological tumor features were obtained from pathological reports. For tumor stage, we used TNM classification (6th edition, WHO, 2002) [24], grouping cases as localized (stage I+II) or invasive (stage III+IV). Tumor site was described according to ICD-O classification (3rd edition, WHO, 2000) [25]; all cases with tumor location coded C18 or C19 were defined as colon cancers, and those coded as C20 were defined as rectal cancers.

The study was reviewed and approved by the Ethics Committee of the National College of Physicians in Romania. All participating subjects were fully informed about the goals and procedures of the study, and written informed consent was obtained and signed prior to enrolment. There was a 98% participation rate for both cases and controls.

Genotyping

Genomic DNA was extracted from peripheral blood using a semi-automated platform for high-quality, high-throughput DNA extraction. The kits (Chemagic DNA Blood10k kit), the equipment (Chemagic Magnetic Separation Module MSM I), and the methods were from Chemagen (Chemagen Biopolymer-Technologie AG). We genotyped rs10795668 on 10p14, rs3802842 on 11q23.1, rs4939827 on 18q21.1, and rs6983267 on 8q24.21. Genotyping of all SNPs and an initial quality check were performed using the Centaurus Assay developed by Nanogen and reagents from EPOCH (Nanogen, Inc.). Genomic DNA extraction, genotyping, and quality control were performed at deCode Genetics, Reykjavik, Iceland.

Statistical analysis

Each SNP was tested for deviation from Hardy-Weinberg equilibrium in the controls by comparing the observed and expected genotypes frequencies with the Chi-square test with one degree of freedom, using a freely available web tool (http://www.oege.org/software/hwe-mr-calc.shtml).

We evaluated the association between the risk of cancer for each SNP and the corresponding risk alleles (as previously reported in the literature) using unconditional multiple logistic regression for additive, dominant, and recessive genetic models. For additive model, subjects were assigned a dummy variable of 0, 1, or 2, representing the number of risk alleles they had for that SNP. For the dominant

model, subjects were coded 1 if they had at least one risk allele and 0 if they had no risk allele. For the recessive model, subjects were assigned 0 if they had two risk alleles and 1 if otherwise. Odds ratios (ORs) were adjusted for gender and age, and the corresponding 95% confidence intervals (CIs) were calculated. Age groups were set as follows: under 60 years, 60-69 years, and over 69 years.

With our sample size, we had an overall power of 80% and a significance level of 0.05 to detect ORs ≥ 1.93 (rs10795668), 1.94 (rs4939827), 1.91 (rs6983267) and an OR ≤ 0.55 (rs6983267), respectively.

We additionally used the Cochrane-Armitage test for trend, which takes into account the individual genotypes rather than just alleles. Subset analysis included stratification according to tumor location (colon or rectum) and TNM stages (I+II or III+IV). All tests were two-sided and considered significant if $P < 0.05$. Statistical analysis was performed using Stata/SE 11 software.

Results

Study population

The characteristics of the CRC cases and controls included in this study are shown in Table I. Ages ranged from 32-94 years for cases, and 19-88 years for controls (medians of 69 and 63 years, respectivelys p=0.0001). Family history of cancer was slightly higher among cases compared to controls (15.68% versus 11.54%) (P=0.474). Among the CRC cases, there were no significant differences in rates of tumor site (57.52% colon cancers versus 42.48% rectal cancers) or tumor stage (55.26% TNM stage I+II versus 44.73% TNM stage III+IV).

Table I. Characteristics of study population

	Cases	Controls	P
	N	N	
Total subjects	153	182	
Gender[a]	N (%)	N (%)	
Women	77 (50.3)	37 (20.3)	0.0001
Men	76 (49.7)	145 (79.7)	
Age at selection	Years	Years	
Mean age ± SD[b]	67.9 ± 11.2	60.8 ± 14.7	0.0001
Median age[c]	69	63	0.0001
Age groups[a]	N (%)	N (%)	
<60 years	29 (18.9)	71 (39.0)	0.0001
60–69 years	48 (31.4)	52 (28.6)	
≥70 years	76 (49.7)	59 (32.4)	
Family history of cancer[a]	N (%)	N (%)	
None	129 (84.3)	161 (88.5)	0.474
1 relative	20 (13.1)	16 (8.8)	
2 relatives	4 (2.6)	5 (2.7)	
Personal cancer history[a]	N (%)	N (%)	
No	140 (91.5)	182 (100)	0.0001
Yes	13 (8.5)	0 (0)	
Site of tumor	N (%)	N (%)	
Colon	88 (57.5)	n/a	
Rectum	65 (42.5)	n/a	
Stage of tumors by TNM	N (%)	N (%)	
Stage I	36 (23.7)	n/a	
Stage II	48 (31.6)	n/a	
Stage III	51 (33.5)	n/a	
Stage IV	17 (11.2)	n/a	

n/a - not applicable; [a]Fisher's exact test; [b]Student's t-test; [c]Wilcoxon rank-sum test

Genotyping data

Descriptions of the genotyped SNPs and the reported risk alleles are presented in Table II. For all four SNPs investigated, the allele frequencies among controls were consistent with the Hardy-Weinberg equilibrium (P≥ 0.01). Table III shows no evidence of a statistically significant association between the investigated SNPs and CRC risk—with one exception—rs4939827 heterozygous and dominant variants, for which we observed a significant inverse association with CRC risk. Stratified analyses for each SNP by tumor location and TNM staging are shown in Tables III-IV, together with ORs and P values.

rs10795668

In our study population, we found no association between CRC and the rs10795668 A allele, which was previously reported to be associated with CRC risk. Tumor location stratified analysis showed a marginally significant risk for rectal cancer compared to colon cancer for AG heterozygous individuals (OR = 2.12; CI = 0.99-4.10; P = 0.052). This was consistent with the results for the dominant model of the A allele in the case of risk of rectal cancer versus colon cancer (OR = 1.68; CI = 0.86-3.26), as well as for dominant model carriers for rectal cancer versus controls (OR = 1.27; CI = 0.69-2.37). We found no evidence for association of rs10795668 with TNM stages of cancer.

rs3802842

The rs3802842 C allele had an association with CRC, with the highest OR found for AC heterozygous (OR = 1.11; CI = 0.68-1.81). However, our stratified analyses revealed that this SNP was associated with a significantly lower risk for rectal cancers compared to colon cancers, having an OR of 0.47 (CI = 0.23-0.94; P = 0.033) for the heterozygous variant, and an OR of 0.44 (CI = 0.22-9.87; P = 0.019) for the dominant model, with a Ptrend of 0.02. These results are supported by the risk estimates for colon cancer versus controls, with an OR of 1.42 (CI = 0.80-2.53) for heterozygous and an OR of 1.40 (CI = 0.80-2.46) for the dominant model, although these estimates were statistically non-significant. We found no association with TNM stages of tumors for rs3802842.

rs4939827

The analysis of rs4939827 T allele carriers showed the above-mentioned unexpected significant inverse association with CRC, with an OR of 0.42 (CI = 0.23-0.76; P = 0.004) for CT heterozygous and an OR of 0.46 (CI = 0.26-0.81; P = 0.008) for dominant variant. When we compared only the

Table II. Description of SNPs selected for genotyping on the basis of genome-wide association studies

SNP	Risk Allele	Region	Reported Gene(s)	Risk Allele Frequency in Controls	P	OR (95% CI)	First Author/Date/ Journal/Study	Initial Sample Size
rs10795668	A	10p14	Intergenic	0.67	3×10^{-13}	1.12 (1.10-1.16)	Tomlinson, March, 2008, Nat Genet	922 cases, 927 controls
rs3802842	C	11q23.1	Intergenic	0.43	6×10^{-10}	1.11 (1.08-1.15)	Tenesa, March 30, 2008, Nat Genet	981 cases, 1,002 controls
rs4939827	T	18q21.1	SMAD7	0.52	8×10^{-28}	1.2 (1.16-1.24)	Tenesa, March, 2008, Nat Genet	981 cases, 1,002 controls
rs6983267	G	8q24.21	Intergenic	0.48	7×10^{-11}	1.24 (1.17-1.33)	Tomlinson, March, 2008, Nat Genet	922 cases, 927 controls

colon cancer cases with controls, we found marked inverse associations for CT heterozygous, with an OR of 0.21 (CI = 0.11-0.43; P = 0.0001); for TT homozygous, with an OR of 0.29 (CI = 0.13-0.65; P = 0.003); and for dominant model, with an OR of 0.24 (CI = 0.12-0.46; P = 0.0001) with a Ptrend of 0.01. At the same time, when comparing rectal cases with colon cases, we found a statistically significant risk for rectal cancers, having a Ptrend = 0.001, for both homozygous (OR = 4.85; CI = 1.76-13.37; P = 0.002) and heterozygous variants (OR = 4.91; CI = 1.92-12.50; P = 0.001), and for the dominant model (OR = 4.88; CI = 2.01-11.85; P = 0.0001). These findings were consistent with the risk of rectal cancer when compared to the controls, with an OR of 1.88 (CI = 0.69-5.13) for homozygous and an OR of 1.48 (CI = 0.60-3.63) for dominant variant, even if these ORs were not statistically significant. We observed statistically significant inverse associations between stages I+II TNM and rs4939827 CT variant (OR = 0.33; CI = 0.16-0.68; P = 0.003) and dominant model (OR = 0.38; CI = 0.19-0.75; P = 0.005), as well as a marginal association between TNM stages III+IV and the heterozygous (OR = 0.20; CI = 0.25-1.12; P = 0.098).

rs6983267

For the G allele of rs6983267, we found an association with CRC, with an OR of 1.38 (CI = 0.70-2.69) for homozygous and an OR of 1.21 (CI = 0.69-2.14) for dominant model. Our stratified analyses by tumor location revealed an increased risk, with statistical significance for rectal versus colon cancers, with an OR of 3.00 (CI = 1.07-8.39; P = 0.036) for homozygous variants and with an OR of 2.61 (CI = 1.06-6.43; P = 0.038) for the dominant model of the G allele. At the same time, we found a greater risk for rectal cancer when compared to controls with an OR of 2.17 (CI = 0.82-5.75) for GG homozygous and an OR of 1.91 (CI = 0.81-4.54) for dominant model, although these risks were statistically non significant. For CRC stages I+II, we observed slightly increased associations for homozygous and for dominant variant (OR = 1.36; CI = 0.60-3.07 and OR = 1.18; CI = 0.59-2.36, respectively). The risk of developing advanced diseases (TNM stages III+IV) was greater for heterozygous (OR = 1.20; CI = 0.56-2.57) and for dominant variant (OR = 1.18; CI = 0.58-2.43). These associations were not statistically significant.

Discussion

Recently, we reported associations of rs10795668, rs3802842, rs6983267, rs4939827, and rs4444235 with CRC risk in a homogenous Romanian population sample, suggesting different patterns of susceptibility for colon and rectal cancers [26]. Here, in an extended case-control series, we pursued these particular findings and we examined four SNP markers that were described in the literature as having tumor site and/or stage-specificity: rs10795668, rs3802842, rs6983267, and rs4939827.

rs4939827

In our analysis, carriers of the T allele had a greater risk of developing rectal cancer (with the most marked association for homozygous), while the C allele conferred susceptibility for colon cancer. This was not surprising, since several other GWAS have shown that common variations at 18q21, defined as rs4939827, rs12953717, and rs4464148, influence CRC risk and show notable site-specific differences [8, 27, 28]. Our results were parallel to those found by Tenesa et al [8], while Curtin et al [27] reported findings in the opposite direction. A possible explanation is that 18q21 contains genes encoding all three types of SMADs required for regulation of TGF-beta signaling, including two activators (SMAD2 and SMAD4) and one inhibitor (SMAD7). Deregulation of TGF-beta signaling is associated with malignant transformation of adenomas driven by CIN, which is more frequent in the rectum [29]. In fact, five (14q22, 15q13, 18q21, 19q13, and 20p12) of the ten independent validated loci associated with CRC are located within proximity of genes involved in TGF-beta signaling, highlighting this tumor-suppressor pathway in the pathogenesis of CRC [5, 6, 8, 9].

rs6983267

Variants at 8q24 have been associated with numerous cancers in addition to CRC, including the prostate [30], breast [31], and urinary bladder [32]. A previous study showed a physical interaction between the rs6983267 region and MYC [33], an important proto-oncogene that is over-expressed in numerous cancers, including CRC. It also was previously demonstrated that SNP rs6983267 at 8q24 enhances Wnt signaling and disrupts the TCF4 transcription factor binding site within 1 Mb of MYC [34]. This deregulated Wnt signaling is predominantly associated

Table III. Association of selected SNPs with colorectal cancer risk, respectively with TNM stage. ORs were adjusted for gender and age group

SNP and risk allele	Geno-type	Ctrs N (%)	CRC Cases, N (%)	Stage I+II cases N (%)	Stage III+IV cases N (%)	CRC cases vs. controls			Stage I+II cancer cases vs. controls			Stage III+IV cancer cases vs. controls		
						OR (95%CI)	P	Ptr	OR (95%CI)	P	Ptr	OR (95%CI)	P	Ptr
rs10795668 A	GG	93 (51.1)	81 (53.6)	44 (52.4)	37 (54.4)		1			1			1	
	AG	70 (38.5)	56 (37.1)	33 (39.3)	24 (35.3)	0.96 (0.59-1.59)	0.888		0.95 (0.52-1.72)	0.869		0.98 (0.32-1.86)	0.947	
	AA	19 (10.4)	14 (9.3)	7 (8.3)	7 (10.3)	0.96 (0.43-2.14)	0.913	0.62	0.87 (0.32-2.34)	0.778	0.75	1.11 (0.40-3.07)	0.843	0.68
	Dom	89 (48.9)	70 (46.4)	39 (46.9)	30 (44.8)	0.96 (0.60-1.54)	0.875		0.93 (0.53-1.63)	0.812		1.00 (0.55-1.83)	0.987	
	Rec	163 (89.6)	137 (90.8)	76 (91.6)	60 (89.5)	1.03 (0.47-2.24)	0.941		1.13 (0.43-2.95)	0.805		0.90 (0.33-2.38)	0.823	
rs3802842 C	AA	101 (55.5)	81 (53.3)	42 (50.0)	38 (56.7)		1			1			1	
	AC	68 (37.4)	65 (42.8)	40 (47.6)	25 (37.3)	1.11 (0.68-1.81)	0.680		1.10 (0.61-1.96)	0.746		0.99 (0.53-1.87)	0.982	
	CC	13 (7.1)	6 (3.9)	2 (2.4)	4 (5.9)	0.68 (0.22-1.08)	0.497	0.91	0.48 (0.09-2.45)	0.376	0.72	0.59 (0.25-3.28)	0.860	0.75
	Dom	81 (44.5)	71 (46.7)	42 (50.0)	29 (43.3)	1.05 (0.65-1.69)	0.840		1.02 (0.58-1.81)	0.929		0.98 (0.53-1.79)	0.942	
	Rec	169 (92.9)	146 (96.1)	82 (97.6)	63 (94.0)	1.54 (0.51-4.64)	0.441		2.18 (0.43-10.9)	0.343		1.12 (0.31-4.03)	0.861	
rs4939827 T	CC	32 (17.7)	42 (27.5)	24 (28.6)	18 (26.5)		1			1			1	
	CT	106 (58.6)	69 (45.1)	35 (41. 7)	33 (48.5)	0.42 (0.23-0.76)	**0.004**		0.33 (0.16-0.68)	**0.003**		0.20 (0.25-1.12)	0.098	
	TT	43 (23.8)	42 (27.5)	25 (29.8)	17 (25.0)	0.56 (0.28-1.10)	0.095	0.45	0.49 (0.22-1.12)	0.090	0.63	0.28 (0.26-1.49)	0.286	0.44
	Dom	149 (82.3)	111 (72.6)	60 (71.4)	50 (73.5)	0.46 (0.26-0.81)	**0.008**		0.38 (0.19-0.75)	**0.005**		0.56 (0.27-1.14)	0.109	
	Rec	138 (76.2)	111 (72.6)	59 (70.2)	51 (75.0)	0.97 (0.57-1.65)	0.900		0.92 (0.49-1.72)	0.787		1.02 (0.51-2.03)	0.960	
rs6983267 G	TT	43 (23.6)	31 (20.5)	16 (19.5)	15 (22.1)		1			1			1	
	GT	92 (50.6)	75 (49.7)	40 (48.8)	34 (50.0)	1.13 (0.62-2.06)	0.691		1.09 (0.52-2.27)	0.826		1.20 (0.56-2.57)	0.636	
	GG	47 (25.8)	45 (29.8)	26 (31.7)	19 (27.9)	1.38 (0.70-2.69)	0.352	0.36	1.36 (0.60-3.07)	0.456	0.29	1.16 (0.49-2.71)	0.736	0.72
	Dom	139 (76.4)	120 (79.5)	66 (80.5)	53 (77.9)	1.21 (0.69-2.14)	0.509		1.18 (0.59-2.36)	0.645		1.18 (0.58-2.43)	0.642	
	Rec	135 (74.2)	106 (70.2)	56 (68.3)	49 (72.1)	0.79 (0.47-1.33)	0.381		0.78 (0.42-1.45)	0.429		0.97 (0.50-1.91)	0.951	

Ptr = P value of Cochrane-Armitage test for trend; Dom: dominant; Rec: recessive model; Ctrs: controls.

with the classic tubular adenoma pathway and CIN cancers, which are more frequent in the rectum [29].

The association between rs6983267 and colon and colorectal cancer in European and European American populations is strongly supported by a recent meta-analysis, which also underlines the fact that allele frequency varies across racial and ethnic groups and may influence the magnitude of associations [35]. The site-specific differences in risk for this variant have been analyzed in a small number of studies. Two independent studies [20, 35] reported significant associations with colon tumors (cases with tumors in the rectosigmoidian junction or rectum were excluded). A recent Swedish study did not detect a significant association with tumor location (colon versus rectum or right versus left colon) [36] Moreover, a separate analysis in cancer-free adenoma cases provided evidence that rs6983267 is associated with an elevated risk of adenoma development [4], especially multiple adenoma [37], pointing to the neoplastic transformation of adenomas in the rectum.

Thus, the association of rs6983267 with rectal cancers,

Table IV. Association of genotypes of the four SNPs with tumor location. ORs were adjusted for gender and age groups.

SNP and risk allele	Genotype	Colon Cancer cases N (%)	Rectal cancer cases N (%)	Colon cancer cases vs. controls			Rectal cancer cases vs. controls			Rectal cancer cases vs. colon cancer cases		
				OR (95%CI)	P	Ptr	OR (%95CI)	P	Ptr	OR (95%CI)	P	Ptr
rs10795668 A	GG	51 (58.6)	30 (46.9)	1			1			1		
	AG	26 (29.9)	30 (46.9)	0.71 (0.38-1.29)	0.261		1.36 (0.72-2.59)	0.346		1.99 (0.99-4.04)	0.054	
	AA	10 (11.5)	4 (6.2)	0.97 (0.39-2.34)	0.942	0.36	0.89 (0.26-2.99)	0.850	0.82	0.78 (0.22-2.75)	0.694	0.33
	AG + AA, vs. GG	36 (41.4)	34 (53.1)	0.76 (0.44-1.33)	0.339		1.27 (0.69-2.37)	0.442		1.68 (0.86-3.26)	0.127	
	AA, vs. AG + GG	77 (88.5)	60 (93.8)	0.90 (0.38-2.15)	0.814		1.30 (0.40-4.20)	0.660		1.72 (0.50-5.87)	0.386	
rs3802842 C	AA	40 (45.4)	41 (64.1)		1			1			1	
	AC	43 (48.9)	22 (34.4)	1.42 (0.80-2.53)	0.232		0.72 (0.38-1.39)	0.332		0.47 (0.23-0.94)	**0.033**	
	CC	5 (5.7)	1 (1.6)	1.25 (0.36-4.32)	0.721	0.19	0.21 (0.02-1.80)	0.154	0.15	0.19 (0.02-1.72)	0.140	**0.02**
	AC + CC, vs. AA	48 (54.5)	23 (35.9)	1.40 (0.80-2.46)	0.238		0.65 (0.34-1.23)	0.187		0.44 (0.22-9.87)	**0.019**	
	CC, vs. AC +AA	83 (94.3)	63 (98.4)	0.95 (0.28-3.17)	0.930		4.25 (0.50-36.12)	0.185		3.73 (0.42-3.21)	0.238	
rs4939827 T	CC	34 (38.6)	8 (12.3)		1			1			1	
	CT	34 (38.6)	35 (53.8)	0.21 (0.11-0.43)	**0.0001**		1.31 (0.52-3.32)	0.564		4.91 (1.92-12.50)	**0.001**	
	TT	20 (22.7)	22 (33.8)	0.29 (0.13-0.65)	**0.003**	**0.01**	1.88 (0.69-5.13)	0.218	0.10	4.85 (1.76-13.37)	**0.002**	**0.001**
	CT + TT, vs. CC	54 (61.4)	57 (87.7)	0.24 (0.12-0.46)	**0.0001**		1.48 (0.60-3.63)	0.390		4.88 (2.01-11.85)	**0.0001**	
	TT, vs. CT + CC	68 (77.3)	43 (66.1)	1.26 (0.64-2.39)	0.520		0.66 (0.34-1.29)	0.226		0.60 (0.29-1.25)	0.173	
rs6983267 G	TT	23 (26. 4)	8 (12.5)		1			1			1	
	GT	41 (47.1)	34 (53.1)	0.91 (0.46-1.80)	0.783		1.78 (0.72-4.39)	0.209		2.41 (0.94-6.20)	0.068	
	GG	23 (26.4)	22 (34.4)	1.01 (0.47-2.19)	0.979	0.82	2.17 (0.82-5.75)	0.118	0.05	3.00 (1.07-8.39)	0.036	0.07
	GT + GG, vs. TT	64 (73.6)	56 (87.5)	0.94 (0.50-1.79)	0.857		1.91 (0.81-4.54)	0.141		2.61 (1.06-6.43)	0.038	
	GG, vs. GT + TT	64 (73. 6)	42 (65.6)	0.93 (0.50-1.74)	0.815		0.71 (0.36-1.39)	0.320		0.63 (0.31-1.31)	0.224	

Ptr = P value of Cochrane-Armitage test for trend

rather than with colon cancers, observed in our study could be due to different population, sample size, or study design.

rs3802842 and rs10795668

It has been demonstrated that susceptibility alleles of rs3802842 and rs10795668 are more prevalent in rectal tumors [7, 8]. Both SNPs show significantly different patterns of associations for different ethnic groups [38]. rs3802842 is located close to POU2AF1, which encodes a transcription factor whose role is yet to be determined. Elucidating the role of rs10795668 is difficult since this SNP is located in an 82-kb LD block that contains no protein coding transcripts [10].

In the present study, we found that the A allele of rs10795668 tended to be associated with rectal cancers, but the differences were not statistically significant. Carrying the C allele at rs3802842 appeared to be associated with a lower risk for rectal tumors, a reverse effect to that previously reported [8, 28, 38].

It is difficult to discuss how these correlations might be interpreted in their biological context, as the exact causative SNP is still not known for all risk loci. However, our findings may be useful for confirming these loci as distinct and etiologically different risk factors in colon and rectum carcinogenesis.

Our study has strengths and limitations. One strength is its well-defined homogeneous sample of a single ethnic group with detailed clinical data. This allowed us to investigate for subset determinants (tumor site and stage). However, it cannot be ruled out that some of our results are falsely positive in the context of the number of tests performed. Although we tested four SNPs, we did not adjust for multiple testing, as this study was set up as a replication to confirm previous findings. A conservative Bonferroni correction for four SNPs would lead to a P value threshold of 0.05/4 = 0.0125.

The lack of significant associations observed, could also be attributed to a limited sample size. In particular, the power would be prohibitively low in subset analyses by stratification according to tumor location. Assuming the ORs of 1.40 for positive associations, and 0.70 for inverse associations at P = 0.05, a CRC case population of about 1560 with an estimated 930 colon cancer cases and 630 rectal cancer cases would be required to reach a power of 80%. This number of 1560 CRC cases, together with a control population of about 1600 would provide ORs $\geq$ 1.23 for positive associations, and $\leq$ 0.82 for negative associations for CRC risk, with a significance level of 0.05 and a power of 80% for all SNPs investigated.

Our data are consistent with other reports on European populations [23, 27, 36] and strengthen previous evidence of true associations. However, the direction of association is variable across different ethnic groups, suggesting genetic heterogeneity or possible gene-environment interaction [39].

The specific pattern of risk for colon and rectal cancers leads to clinical and public health implications. In the shorter term, the potential to refine genetic risk stratification within populations and phenotypic subgroups could result in the adapting of surveillance to individual-predicted level of risk. In the future, understanding the functional role of these SNPs on the genetic pathways could lead to targeted therapeutic strategies and prognosis.

Conclusions

The present data suggest that, in an independent Romanian sample, common genetic variations at rs10795668, rs3802842, rs4939827, and rs6983267 display different patterns of association with colon and rectal cancer risk, supporting the hypothesis of different carcinogenic pathways with a distinct molecular biology.

Our results should be considered cautiously, but may be taken into account in future, larger, epidemiological studies.

Acknowledgements

This work was supported by CNCSIS-UEFISCSU, project PNII-IDEI, code ID 1184/2008 and by deCode Genetics, Reykjavik, Iceland (Collaboration Agreement, April 8, 2008). We gratefully acknowledge the contribution of Thorunn Rafnar (deCode Genetics, Reykjavik, Iceland) for the genotyping data acquisition and her valuable help with the manuscript.

Conflict of interest statement

The authors confirm that there are no conflicts of interest.

References

1. Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002. *Cancer incidence, mortality and prevalence worldwide*. IARC CancerBase no. 5, version 2.0, IARC Press, Lyon; 2004.
2. Liljegren A, Lindblom A, Rotstein S, Nilsson B, Jaramillo E. Prevalence and incidence of hyperplastic polyps and adenomas in familial colorectal cancer: correlations between the two types of colon polyps. Gut 2003; 52: 1140-1147.
3. Lindgren G, Liljegren A, Jaramillo E, Rubio C, Lindblom A. Adenoma prevalence and cancer risk in familial non-polyposis colorectal cancer. Gut 2002; 50: 228-234.
4. Tomlinson I, Webb E, Carvajal-Carmona L, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet 2007; 39: 984-988.
5. Jaeger E, Webb E, Howarth K, et al. Common genetic variants at CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. Nat Genet 2008; 40: 26-28.
6. Broderick P, Carvajal-Carmona L, Pittman AM, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. Nat Genet 2007; 39: 1315-1317.
7. Tomlinson IP, Webb E, Carvajal-Carmona L, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nat Genet 2008; 40: 623-630.
8. Tenesa A, Farrington SM, Prendergast JG, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat Genet 2008; 40: 631-637.
9. Houlston RS, Webb E, Broderick P, et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. Nat Genet 2008; 40: 1426-1435.
10. Picelli S, Von Holst S, Wessendorf P. The continuing search for predisposing colorectal cancer variants. Cancer Genomics Proteomics 2009; 6: 305-316.
11. Tenesa A, Dunlop MG. New insights into the aetiology of colorectal cancer from genome-wide association studies. Nat Rev Genet 2009; 10: 353-358.
12. Li FY, Lai MD. Colorectal cancer, one entity or three. J Zhejiang Univ Sci B 2009; 10: 219-229.
13. Pinheiro M, Ahlquist T, Danielsen SA, et al. Colorectal carcinomas with microsatellite instability display a different pattern of target gene mutations according to large bowel site of origin. BMC Cancer 2010; 10: 587.
14. Osian G, Procopciuc L, Vlad L, et al. NAT2 polymorphisms and sporadic colorectal cancer survival. J Gastrointestin Liver Dis 2010; 19: 361-368.

15. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med 2004; 10: 789-799.
16. Fernebro E, Halvarsson B, Baldetorp B, Nilbert M. Predominance of CIN versus MSI in the development of rectal cancer at young age. BMC Cancer 2002; 2: 25.
17. Kapiteijn E, Liefers GJ, Los LC, et al. Mechanisms of oncogenesis in colon versus rectal cancer. J Pathol 2001; 195: 171-178.
18. Gervaz P, Bouzourene H, Cerottini JP, et al. Dukes B colorectal cancer: distinct genetic categories and clinical outcome based on proximal or distal tumor location. Dis Colon Rectum 2001; 44: 364-372.
19. Colombino M, Cossu A, Manca A, et al. Prevalence and prognostic role of microsatellite instability in patients with rectal carcinoma. Ann Oncol 2002; 13: 1447-1453.
20. Cicek MS, Slager SL, Achenbach SJ, et al. Functional and clinical significance of variants localized to 8q24 in colon cancer. Cancer Epidemiol Biomarkers Prev 2009; 18: 2492-2500.
21. Haiman CA, Le Marchand L, Yamamato J, et al. A common genetic risk factor for colorectal and prostate cancer. Nat Genet 2007; 39: 954-956.
22. Tuupanen S, Niittymaki I, Nousiainen K, et al. Allelic imbalance at rs6983267 suggests selection of the risk allele in somatic colorectal tumor evolution. Cancer Res 2008; 68: 14-17.
23. Ghazi S, von Holst S, Picelli S, et al. Colorectal cancer susceptibility loci in a population-based study. Am J Pathol 2010; 177: 2688–2693.
24. TNM Classification of Malignant Tumors. 6th ed, New York: Wiley-Liss, 2002, 72-76.
25. International Classification of Diseases for Oncology. 3rd ed, Geneva: World Health Organization, 2000, 49.
26. Mates IN, Csiki I, Mates D, et al. Association of common genetic variants with colorectal cancer in a Romanian sample. Chirurgia (Bucur) 2010; 105: 749-757.
27. Curtin K, Lin WY, George R, et al. Meta association of colorectal cancers confirms risk alleles at 8q24 and 18q21. Cancer Epidemiol Biomarkers Prev 2009; 18: 616-621.
28. Pittman AM, Webb E, Carvajal-Carmona L, et al. Refinement of the basis and impact of common 11q23.1 variation to the risk of developing colorectal cancer. Hum Mol Genet 2008; 17: 3720-3727.
29. Pritchard CC, Grady WM. Colorectal cancer molecular biology moves into clinical practice. Gut 2011; 60: 116-129.
30. Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet 2006; 38: 652-658.
31. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007; 447: 1087-1093.
32. Kiemeney LA, Thorlacius S, Sulem P, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat Genet 2008; 40: 1307-1312.
33. Pomerantz MM, Ahmadiyeh N, Jia L, et al. The 8q24 cancer risk variants rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet 2009; 41: 882-884.
34. Tuupanen S, Turunen M, Lehtonen R, et al. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat Genet 2009; 41: 885-890.
35. Hutter CM, Slattery ML, Duggan DJ, et al. Characterization of the association between 8q24 and colon cancer: gene-environment exploration and meta-analysis. BMC Cancer 2010; 10: 670.
36. von Holst S, Picelli S, Edler D, et al. Association studies on 11 published colorectal cancer risk loci. Br J Cancer 2010; 103: 575-580.
37. Berndt SI, Potter JD, Hazra A, et al. Pooled analysis of genetic variation at chromosome 8q24 and colorectal neoplasia risk. Hum Mol Genet 2008; 17: 2665-2672.
38. He J, Wilkens LR, Stram DO, et al. Generalizability and epidemiologic characterization of eleven colorectal cancer GWAS hits in multiple populations. Cancer Epidemiol Biomarkers Prev 2011; 20: 70-81.
39. Sebastiani P, Timofeev N, Dworkis DA, Perls TT, Steinberg MH. Genome-wide association studies and the genetic dissection of complex traits. Am J Hematol 2009; 84: 504-515.