Referate generale

Chirurgia (2010) 105: 749-757
Nr. 6, Noiembrie - Decembrie

Association of common genetic variants with colorectal cancer risk in a Romanian sample

I.N. Mateș[1], I. Csiki[2], D. Mateș[2], V. Constantinescu[2], P. Badea[2], D. Dinu[1], A. Constantin[1], S. Constantinoiu[1]

[1]St. Mary Clinic of General and Esophageal Surgery, Carol Davila Universtity of Medicine and Pharmacy, Bucharest
[2]National Institute of Public Health, Bucharest

Rezumat

Variante genetice comune asociate riscului de cancer colorectal pe un lot din România

Recent, mai multe studii de asociere genomică au identificat și validat locusuri unde variantele genetice comune influențează riscul de cancer colorectal. Am testat asocierea dintre opt SNPuri și cancerul colorectal, într-un studiu caz-control efectuat pe un lot din România. Am genotipat rs10795668, rs16892766 rs3802842, rs4444235, rs4779584, rs4939827, rs6983267, rs9929218 și am testat asocierea lor statistică cu afecțiunea. Cinci SNPuri (rs6983267, rs4939827, rs3802842, rs4444235, rs10795668) s-au asociat cu riscul de cancer colic și rectal. Pentru trei asocierea a fost semnificativă statistic: alelele de risc ale rs6983267 și rs4939827 au fost asociate semnificativ cancerului rectal (p=0.031 și p=0.004 pentru homozigoți, p=0.002 și p=0.005 pentru heterozigoți). Pentru rs3802842 riscul de cancer colic a fost mai mare decât riscul de cancer rectal (OR=2.26; CI=1.04-5.88, p=0.040), pentru modelul dominant. rs4444235 a confirmat riscul atât pentru homozigoți cât și pentru heterozigoți, având cel mai mare OR de 1.49(CI=0.61-3.61) pentru heterozigoți. rs10795668 a fost asociat cu un risc mai mare de cancer rectal față de martori (OR=1.46; CI=0.66-3.21) și un risc mai mare de cancer rectal față de colon (OR=2.19; CI=0.87-5.55). Acesta este primul studiu din România care confirmă asocierea dintre cancerul colorectal și cinci din opt SNPuri investigate, subliniind necesitatea replicării rezultatelor pe loturi mai mari.

Cuvinte cheie: cancer colorectal, SNP, studiu de asociere

Abstract

Recently, several genome-wide association studies identified and validated loci at which common genetic variants influence the risk of colorectal cancer. We aimed to test the association between eight SNPs and colorectal cancer in a Romanian case-control sample. We genotyped rs10795668, rs16892766, rs3802842, rs4444235, rs4779584, rs4939827, rs6983267, and rs9929218 and we statistically tested the association with the disease. Five SNPs (rs6983267, rs4939827, rs3802842, rs4444235, rs10795668) showed an association with colon and rectal cancer. Three of them proved to be statistically significant: rs6983267 and rs4939827 risk alleles were significantly associated with rectal cancers (p=0.031 and p=0.004 for homozygous, p=0.002 and p=0.005 for heterozygous). For rs3802842 we found a greater risk for colon than rectal cancer with an OR of 2.26(CI=1.04-5.88, p=0.040) for the dominant model. The rs4444235 confirmed the risk for both homozygous and heterozygous carriers, with the greatest ORs of 1.49(CI=0.61-3.61) for heterozygote. For rs10795668 we found an increased risk for rectum cancer vs. controls with an OR of 1.46(CI=0.66-3.21), and for rectum cancer vs. colon cancer (OR=2.19; CI=0.87-5.55). This is the first Romanian study that confirms previously-identified associations with colorectal cancer risk for five out of eight

Corresponding author: Dr. Ioan Nicolae Mates
St. Mary Clinic of General and Esophageal Surgery
37-39 Ion Mihalache blv, sector 1,
78218, Bucharest, Romania
E-mail: ioan.mates@gmail.com

SNPs investigated and underlines the necessity of extensive replication using larger samples.

Key words: colorectal cancer, SNPs, association study

Introduction

Colorectal cancer (CRC) is the third most common malignancy and the fourth cause of cancer mortality worldwide. Over 1 million new cases are diagnosed each year and incidence seems set to rise (1). The overall mortality rate is about 50%, the single most important survival factor being the extent of tumour progression at diagnosis.

In Romania, CRC is the second common cancer among men (standardized incidence rate 40.69/100,000) and the third among women (25.10/100,000). CRC is the second cause of cancer mortality in men (after lung cancer), and the third cause of cancer mortality in women (after breast and cervix cancers), with a standardized mortality rate of 23.48/100,000 and 14.55/100,000 respectively (2,3). In fact, the deaths caused by CRC represent 11.76% of all deaths caused by cancer in men and 9.71% in women.

In 1955-1959 mortality rate was 4.65/100,000 in males and 4.57/100,000 in females. In 1990-1992 the mortality increased to 10.10/100,000 and 7.40/100,000 respectively. After 10 years the rate doubled to 19.20/100,000 (4).

The constant rise of the mortality rate, compared to the decreasing slope in other European countries, raises many public health issues, mainly prevention of the disease (5).

CRC is a complex trait influenced by genetic and environmental factors and their interactions. About 75% of patients with colorectal cancer have sporadic disease, with no apparent evidence of having inherited the disorder. Heredity is responsible for approximately one third of our susceptibility to colorectal cancer and the causative germline mutations account for less than 6% of all CRC cases (6,7).

Genetic mutations in major genes have been identified as the cause of inherited cancer risk in some colon cancer –prone families. The functions of the major colon cancer genes have been reasonably well characterized over the past decade (6,8). Three proposed classes of colon cancer genes are tumour suppressor genes (APC, AXIN2, TP53, STK11, PTEN, BMPR1A, SMAD4), oncogenes (KIT, PDGFRA), and stability genes (MLH1, MYH, BLM).

It is currently believed that many of the moderate and low-risk cases are influenced by low penetrance genes or gene combinations. These types of genetic variations are often referred to as polymorphisms. Most loci that are polymorphic have no influence on disease risk or human traits (benign polymorphisms), while those that are associated with a difference in risk of disease or a human trait (however subtle) are sometimes termed disease-associated polymorphisms or functionally relevant polymorphisms.

In general, low-penetrance variants have been identified in one of two manners. Earlier studies focused on candidate genes chosen because of biologic relevance to colon cancer pathogenesis (association studies). Candidate alleles that have been shown to associate with modest increased frequencies of colon cancer include heterozygous BLM^{Ash}, the GH1 1663 T→A polymorphism and the APC I1307K polymorphism (9, 10).

More recent approaches have been conducted with relatively large, unselected series of CRC patients that have been evaluated for patterns of polymorphisms in candidate and anonymous genes spread throughout the genome. Population genetic theory predicts that distribution of allelic effects influencing complex diseases is L shaped with a small number of variants having a large phenotypic effect and a large number of variants having an individually small effect (11). The most common type of variation is a single nucleotide base substitution known as the single-nucleotide polymorphism (SNP), which by definition is observed in at least 1% of the population. Most common variants of SNPs namely those with a minor allele frequency (MAF) > 5% are common to all population, but the distribution of allele frequency can vary largely across the globe (12). The International HapMap project launched in 2003 made possible the characterization of the common SNPs and the description of the block structure of the human genome that could be used to identify the tagSNPs (13,14). Based on the results of HapMap project, the genome wide association studies (GWAS) emerged as a powerful tool to identify susceptibility loci with low effect in unrelated subjects with specific cancers and related outcomes. Typically, a GWAS uses a case-control design: cases are diseased subjects and controls should be disease-free. The choice of the SNPs to be typed can vary: either tag SNPs chosen to capture substantial variability of the common SNPs, or randomly selected SNPs, or a combination of both strategies. The major genotyping platforms that are currently available offer similar coverage of common variants in the HapMap ranging between 500,000 and 1,000,000 SNPs per sample. The core of the analysis is a statistical test conducted for each SNP in the array with the objective to select those SNPs showing significant association with the trait per phenotype studied. Replication of results is critical in a separate comparable set of studies; the value of the replication is to guard against the risk of false positives observed with common alleles with low effect sizes (12).

The GWAS opened the door to the discovery of several common genetic variants involved in the aetiology of CRC. Until mid 2007 no common variants contributing to the CRC risk had been identified and replicated. Ten common low penetrance variants have since been identified and replicated in UK (15). However, the effect size of these genetic variants was modest with an Odds Ratio (OR) of 1.2. In fact, a meta-analysis of all UK GWAS data leads the authors to the followings conclusion (16):

- the loci readily detectable through current GWAS are associated with modest effects genotypic risks of approximately 1.2;
- the number of common variants explaining more

than 1% of inherited risk is very low;
- only a small proportion of heritability of any cancer can be explained by the currently identified loci;
- of the common risk loci identified thus far, no significant epistatic effects were observed.

Because few of the observed associations seem to be due to correlation with common coding variants and many of the loci map to regions lacking genes of protein-coding transcripts, it seems likely that much of the common variation in cancer risk is mediated through sequence changes influencing gene expression.

This study is a collaborative on-going investigation with deCODE Genetics, Iceland, aimed to test previously identified and validated SNPs for association with CRC in a Romanian case-control series, to evaluate if these differed by several phenotypic traits, and to pilot the association with CRC in a nested Romanian case-control group enrolled between 2007 and 2009.

Materials and methods

Colorectal cases were identified among patients hospitalized at St. Mary Clinic of General and Esophageal Surgery Clinic of the Carol Davila University of Medicine and Pharmacy, Bucharest. Cases were ascertained on the basis of their histological report. Controls were hospitalized in the same clinic and selected to be cancer-free subjects.

All eligible subjects were invited to participate and were fully informed about the goals and the procedures of the study. A written informed consent was obtained and signed prior to enrolment. The study was approved by the Bioethical Committee of the Romanian College of Physicians.

We included 92 cases with CRC histological confirmed and 96 cancer-free controls, hospitalized between September 2007 and December 2009. Information was collected by trained interviewers via direct interview with each subject, using a standardized structured questionnaire designed to collect demographic and potential risk factors data. All participants were individuals of self-reported Romanian nationality. Personal and family cancer self-reported history was recorded during interview in order to test the Amsterdam II criteria for hereditary CRC. Clinical information related to CRC, including age at onset, was extracted from the hospital records for each case. Each patient's tumour was staged using the TNM staging and graded using the ICD-O (17,18). A whole-blood sample of 2 x 9 ml was collected for each subject. The 9 ml EDTA vacuteiners were bar-code labeled and stored at -80° Celsius before shipment on dry ice to deCODE Genetics laboratories (Reykjavik, Iceland) for genotyping.

Loci and SNPs

We selected eight SNPs: six replicated in large GWAS prior to 2008 and two validated by Cogent study (16), and the corresponding risk alleles *(Table 1)*. Thus, we genotyped

Table 1. *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-value	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
rs16892766	A	8q23.3	EIF3H	0.07	3×10^{-18}	1.27 (1.20-1.34)	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
rs4444235	C	14q22.2	BMP4	0.46	8×10^{-10}	1.11 (1.08-1.15)	COGENT Study, Nov. 2008 *Nat Genet*	1,902 cases, 1,929 controls
rs4779584	T	15q13.3	Intergenic	0.19	5×10^{-7}	1.23 (1.14-1.34)	Tomlinson March, 2008 *Nat Genet*	922 cases, 927 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
rs9929218	A	16q22.1	CDH1	0.29	1×10^{-8}	1.1 (1.06-1.12)	COGENT Study, Nov. 2008 *Nat Genet*	1,902 cases, 1,929 controls

rs10795668 on 10p14, rs16892766 on 8q23.3, rs3802842 on 11q23.1, rs4444235 on 14q22.2, rs4779584 on 15q13.3, rs4939827 for 18q21.1, rs6983267 for 8q24.21, and rs9929218 on 16q22.1. The selected SNPs showed a modest, but significant effect size on CRC risk replicated in independent studies on different populations (19).

Genotyping

Genomic DNA extraction from peripheral blood was done 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). Genotyping of all eight SNPs and a first quality check were done using 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

Genotypes frequencies for all eight SNP markers in controls and cases were assessed for deviation from Hardy-Weinberg equilibrium (HWE) using Chi square test with one grade of freedom using a freely available web tool (http://www.oege.org/software/hwe-mr-calc.shtml).

Allelic frequencies of SNPs in cases and controls were compared using Chi square tests (non-risk allele 1 vs. risk allele 2). Analyses were performed under different types of genetic models, i.e. the comparison of homozygote versus heterozygote (genotype 11 vs. 12), comparison of homozygote (genotype 11 vs. 22), the dominant model (11 vs. 12+22) or recessive model (11+12 vs. 22).

Odds ratios (OR) together with 95% confidence interval were calculated using a multiple logistic regression model to examine the association between each selected SNP and risk of cancer. ORs were adjusted for gender and age groups of subjects when analyzing the genotypes distribution between cases (CRC, colon or rectum) and controls. When the analysis was aimed at comparing the genotypes distribution between cases of colon and rectum cancers, the ORs were adjusted for gender and age group at onset. Both age groups were set as follows: under 60 years, 60-69 years and over 69 years. If the OR-value for a specified SNP had fluctuating values about 1.0 for the analyzed genotypes, we have considered that there is no association between this SNP and cancer risk. If the OR for the analyzed genotypes had values consistently above or below 1.0, we considered that this indicates the presence of an association between the SNP and cancer. This association can be statistically significant or non-significant, depending on P-value.

In addition, Cochrane-Armitage test for trend, which takes into account the individual genotypes rather than just alleles, was performed.

Continuous variables were compared with Student t test, while nonparametric variables were compared with Chi-test. A level of $p < 0.05$ was used to indicate statistical significance. Statistical analysis was performed using the Stata 7.0 software.

Results

Study population

Descriptions of the CRC cases and controls included in the study are shown in the *Tables 2* and *3*. No significant differences in descriptors between cases and controls were observed. The median age of CRC cases at recruitment was 68.5 years (32 to 85 years), while the median age at onset was

Table 2. *Description of cases and controls*

	Controls	Cases
Total subjects, N	96	92
Gender		
Women, N(%)	37 (38.54 %)	43 (46.74%)
Men, N(%)	59 (61.46 %)	49 (53.26 %)
Age at selection		
Mean age ± SD (years)	62.37 ± 13.87	65.78 ± 12.13
Median age (years)	64	68.5
Age groups, N(%)		
< 60 years	34 (35.42%)	25 (27.17%)
60 – 69 years	27 (28.13%)	26 (28.26%)
≥ 70 years	35 (36.46%)	41 (44.57%)
Family history of cancer, N(%)		
None	89 (92.71%)	79 (85.87%)
1 relative	4 (4.17%)	10 (10.87%)
2 relatives	3 (3.13%)	3 (3.26%)
Personal cancer history, N(%)		
No	96(100%)	83 (90.22%)
Yes	0(0.0%)	9 (9.78%)

Table 3. *Characteristics of cases (N=92)*

Site of tumor, N(%)	
Proximal colon	27 (29.35%)
Distal colon	23 (25.00%)
Rectal	39 (42.39%)
Exact location not specified	3 (3.26%)
Age at onset	
Mean age ± SD (years)	64.94 ± 12.04
Median age (years)	67
Age groups, N(%)	
< 60 years	25 (27.17%)
60 – 69 years	26 (28.26%)
≥ 70 years	41 (44.57%)
Stage, TNM, N(%)	
Stage I	17 (18.48%)
Stage II	31 (33.70%)
Stage III	33 (35.87%)
Stage IV	11 (11.96%)

Table 4. *Risk allele frequencies in patients and controls*

SNP	Risk allele	Controls (N=96)	Cases (N=92)	Ratio of frequencies of risk allele (cases/controls)	P (cases vs. controls)
rs10795668	A	0.286	0.280	0.978	0.924
rs16892766	A	0.927	0.934	1.008	0.835
rs3802842	C	0.239	0.266	1.111	0.674
rs4444235	C	0.479	0.521	1.088	0.559
rs4779584	T	0.218	0.179	0.820	0.675
rs4939827	T	0.542	0.494	0.912	0.652
rs6983267	G	0.546	0.555	1.014	0.911
rs9929218	A	0.369	0.298	0.808	0.303

67 years (30 to 84 years). The median age of controls was 64 years (20 to 88 years). Cases had a higher proportion of family history of cancer compared to the controls (14.03 % vs. 7.30%). None of the cases meets the Amsterdam II criteria. The majority of the tumours were located in the colon (54.35%). 52.18% patients had less advanced tumours (TNM stage I and II). The predominant morphologic feature was M 8140/3, G2 (moderately differentiated adenocarcinoma).

Genotyping data

Data on risk alleles distributions in CRC cases and controls are presented in *Table 4*. No significant differences between cases and controls or between non-risk and risk alleles were observed. The risk (minor) allele frequencies for all SNPs were similar with those reported by the dbSNP for the CEU population (http://www.ncbi.nlm.nih.gov/projects/SNP/).

Genotypes frequencies in CRC cases and controls as well as ORs and P-values for all eight SPNs analyzed are shown in the *Table 5*. Among controls, genotype frequencies for all the analyzed polymorphisms were in agreement with Hardy-Weinberg expectations. Stratified distribution of genotypes by tumour location, colon and rectum vs. control, as well as rectum vs. colon ORs and P-values are shown in the Table 6. Out of eight SNPs analyzed, three associations proved to be statistically significant.

We found no association with colorectal cancer in our population for the rs10795668 AA genotype. However, our stratified analyses revealed an increased risk for rectum cancer vs. controls with the highest OR of 1.46 (CI=0.66-3.21) for the homozygous state, and for rectum cancer vs. colon cancer with a 2-fold increased risk (OR=2.19; CI=0.87-5.55) for the same homozygous state.

No association with CRC was found for rs16892766 AA genotype.

Regarding variant rs3802842 CC genotype, we found an increased risk for CRC (highest OR=1.19; CI=0.65-2.18, homozygous state) and for colon cancer vs. controls (greatest OR=1.67; CI=0.39-7.01, heterozygous state). We found a greater risk for colon than rectal cancer with a statistically significant ptrend of 0.03 associated with OR of 2.26(CI=1.04-5.88) for the dominant model having a P-value=0.040.

The SNP rs4444235 CC genotype confirmed the risk for CRC for both homozygous and heterozygous carriers, with the greatest ORs of 1.49(CI=0.61-3.61, heterozygote CRC), of 1.55(CI=0.58-4.49, homozygote rectum cancer), of 1.23(CI=0.43-3.57, heterozygote colon cancer), as well as of 1.59(CI=0.49-5.16, homozygote rectum vs. colon cancer).

The SNPs rs4779584 TT genotype and rs9929218 AA genotype, in our population, appeared rather having a protective effect for CRC, as well as for colon or rectum cancers, for both homozygous and heterozygous states.

Concerning the SNP rs4939827 CC genotype, this SNP showed a statistic significant protective effect against CRC with OR=0.35(CI=0.16-0.74, p=0.006, homozygous state) and OR=0.42(CI=0.21-0.86, p=0.018, for dominant variant). This protective effect against colon cancer was confirmed for both homozygous and heterozygous states with the most marked protective OR of 0.17(CI=0.07-0.40, p=0.0001, heterozygous carrier) together with a statistically significant ptrend of 0.01. At the same time, the stratified analysis showed an increased and statistically significant risk with a ptrend=0.001 for rectum cancer vs. controls having the highest OR of 2.01(CI=0.60-6.81, heterozygous state). When compared the rectal cancer cases with colon ones, we found an important and statistically significant risk having a ptrend=0.001 for rs4939827 CC genotype for rectal cancers, for both homozygous (OR=5.61; CI=1.70-18.40; p=0.004) and heterozygous variants (OR=6.16; CI=1.75-21.68, p=0.005), including the dominant model (OR=5.63; CI=1.93-17.60; p=0.002).

For SNP rs6983267 GG genotype, no association was found with colorectal cancer. However, our stratified analyses revealed an increased but statistically non-significant risk for rectum cancer vs. controls with the highest OR of 2.57(CI=0.72-9.12) for the heterozygous state. This SNP rs6983267 showed an increased risk, with statistical significance for rectum vs. colon cancers, for both homozygous and heterozygous variants, including the dominant model, having the OR of 4.42(CI=1.22-14.08, p=0.031) for the homozygous state and the OR of 5.63(CI=1.93-17.60, p=0.002) for the heterozygote state.

Table 5. *Associations of genotypes of the eight selected SNPs with colorectal cancer risk. ORs were adjusted for gender and age groups*

SNP and risk allele	Genotype	Controls, N(%)	Cases, N(%)	Ptrend	OR(95%CI)	P
rs10795668	GG	49(51.04%)	48(52.75%)		1.00	
A	AG	39(40.63%)	35(38.46%)		0.97(0.52-1.79)	0.922
	AA	8(8.33%)	8(8.79%)	0.86	1.01(0.35-2.95)	0.981
	AG + AA, vs. GG (Dom)	47(48.96%)	43(47.25%)		0.98(0.55-1.75)	0.939
	AA, vs. AG + GG (Rec)	88(91.67%)	83(91.21%)		1.02(0.36-2.89)	0.961
rs16892766	CC	1(1.04%)	0(0.00%)		1.00	
A	AC	12(12.50%)	12(13.04%)		-	
	AA	83(86.46%)	80(86.96%)	0.90	-	
	AC + AA, vs. CC (Dom)	95(98.96%)	92(100.0%)		-	
	AA, vs. AC + CC (Rec)	13(13.54%)	12(13.04%)		1.04(0.45-2.44)	0.923
rs3802842	AA	55(57.29%)	48(52.17%)		1.00	
C	AC	36(37.50%)	39(42.39%)		1.19(0.65-2.18)	0.565
	CC	5(5.21%)	5(5.43%)	0.50	1.01(0.27-3.78)	0.986
	AC + CC, vs. AA (Dom)	41(42.71%)	44(47.83%)		1.17(0.65-2.10)	0.596
	CC, vs. AC +AA (Rec)	91(94.79%)	87(94.57%)		0.93(0.26-3.40)	0.919
rs4444235						
C	TT	23(23.96%)	18(19.57%)		1.00	
	CT	54(56.25%)	52(56.52%)		1.25(0.60-2.59)	0.548
	CC	19(19.79%)	22(23.91%)	0.38	1.49(0.61-3.61)	0.383
	CT + CC, VS. TT (Dom)	73(76.04%)	74(80.43%)		1.31(0.65-2.64)	0.453
	CC, vs. CT + TT (Rec)	77(80.21%)	70(76.09%)		1.26(0.62-2.56)	0.521
rs4779584						
T	CC	59(61.46%)	60(65.22%)		1.00	
	CT	32(33.33%)	31(33.70%)		0.97(0.52-1.80)	0.925
	TT	5(5.21%)	1(1.09%)	0.47	0.21(0.02-1.87)	0.163
	CT + TT, vs. CC (Dom)	37(38.54%)	32(34.78%)		0.87(0.48-1.59)	0.659
	TT, vs. CT + CC (Rec)	91(94.79%)	91(98.91%)		0.21(0.02-1.87)	0.163
rs4939827						
T	CC	15(15.79%)	28(30.43%)		1.00	
	CT	57(60.00%)	37(40.22%)		0.35(0.16-0.74)	0.006
	TT	23(24.21%)	27(29.35%)	0.38	0.60(0.26-1.40)	0.240
	CT + TT, vs. CC (Dom)	80(84.21%)	64(69.57%)		0.42(0.21-0.86)	0.018
	TT, vs. CT + CC (Rec)	72(75.79%)	65(70.65%)		1.24(0.64-2.39)	0.520
rs6983267						
G	TT	19(19.79%)	21(23.08%)		1.00	
	GT	49(51.04%)	39(42.86%)		0.74(0.35-1.58)	0.443
	GG	28(29.17%)	31(34.07%)	0.84	1.03(0.46-2.34)	0.935
	GT + GG, vs. TT (Dom)	77(80.21%)	70(76.92%)		0.85(0.42-1.72)	0.648
	GG, vs. GT + TT (Rec)	68(70.83%)	60(65.93%)		1.27(0.67-2.37)	0.461
rs9929218						
A	GG	36(37.50%)	45(48.91%)		1.00	
	AG	49(51.04%)	39(42.39%)		0.64(0.34-1.18)	0.149
	AA	11(11.46%)	8(8.70%)	0.12	0.61(0.21-1.69)	0.338
	AG + AA, vs. GG (Dom)	60(62.50%)	47(51.09%)		0.63(0.35-1.14)	0.124
	AA, vs. AG + GG (Rec)	85(88.54%)	84(91.30%)		0.77(0.29-2.03)	0.594

Dom = dominant model, Rec = recessive model

Discussion

For the last three years, multicentric GWAS unraveled the complex genetic architecture behind the sporadic CRC. To date, ten susceptibility loci were identified as been associated with CRC and the results were replicated in various population.

Eight SNPs (rs10795668, rs16892766, rs3802842, rs4444235, rs4779584, rs9929218, rs4939827, rs6983267) mapping seven different region were selected for genotyping

Table 6. *Association of genotypes of the eight SNPs with tumour location. Ptrends and ORs of colon and rectal cancer cases versus controls*

SNP and risk allele	Genotype	Colon Cancer N (%)	Rectal cancer N (%)	Ptrend	Colon vs. control OR*(95%CI)	P	Ptrend	Rectal vs. control OR*(95%CI)	P	Ptrend	Rectal vs. colon OR**(95%CI)	P
rs10795668	GG	30(61.22%)	17(43.59%)		1.00			1.00			1.00	
A	AG	15(30.61%)	19(48.72%)		0.68(0.32-1.47)	0.330		1.46(0.66-3.21)	0.346		2.19(0.87-5.55)	0.096
	AA	4(8.16%)	3(7.69%)	0.30	0.76(0.21-2.78)	0.677	0.51	1.18(0.28-5.07)	0.820	0.15	1.39(0.27-7.14)	0.692
	AG + AA, vs. GG (Dom)	19(8.16%)	22(56.41%)		0.70(0.34-1.42)	0.323		1.42(0.66-3.02)	0.370		2.03(0.84-4.87)	0.115
	AA, vs. AG + GG (Rec)	45(91.84%)	36(92.31%)		0.87(0.24-3.10)	0.833		0.99(0.24-3.99)	0.985		0.99(0.20-4.87)	0.998
rs16892766	CC	0(0.00%)	0(0.00%)		1.00			1.00			1.00	
A	AC	8(16.00%)	4(10.26%)		-	-		-	-		-	-
	AA	42(84.00%)	35(89.74%)	0.71	-	-	0.59	-	-	0.43	1.58(0.43-5.79)	0.488
	AC + AA, vs. CC (Dom)	50(100.0%)	39(100.0%)		-	-		-	-		-	-
	AA, vs. AC + CC (Rec)	8(16.00%)	4(10.26%)		0.79(0.29-2.09)	0.636		1.37(0.41-4.54)	0.603		0.63(0.17-2.31)	0.488
rs3802842	AA	22(44.00%)	26(66.67%)		1.00			1.00			1.00	
C	AC	24(48.00%)	12(30.77%)		1.54(0.74-3.18)	0.248		0.69(0.30-1.56)	0.369		0.43(0.17-1.05)	0.066
	CC	4(8.00%)	1(2.56%)	0.12	1.67(0.39-7.01)	0.490	0.29	0.38(0.04-3.54)	0.400	0.03	0.22(0.02-2.13)	0.192
	AC + CC, vs. AA (Dom)	28(56.00%)	13(33.33%)		1.55(0.77-3.13)	0.222		0.65(0.29-1.43)	0.288		0.39(0.17-0.96)	0.040
	CC, vs. AC +AA (Rec)	46(92.00%)	38(97.44%)		1.35(0.33-5.44)	0.675		0.44(0.05-3.99)	0.469		0.31(0.03-2.94)	0.309
rs4444235	TT	11(22.00%)	6(15.38%)		1.00			1.00			1.00	
C	CT	28(56.00%)	24(61.54%)		1.09(0.46-2.58)	0.847		1.61(0.58-4.49)	0.364		1.59(0.49-5.16)	0.441
	CC	11(22.00%)	9(23.08%)	0.72	1.23(0.43-3.57)	0.701	0.34	1.55(0.45-5.35)	0.489	0.58	1.54(0.39-6.00)	0.536
	CT + CC, vs TT (Dom)	39(78.00%)	33(84.62%)		1.12(0.49-2.58)	0.783		1.07(0.42-2.73)	0.883		1.57(0.50-4.91)	0.435
	CC, vs. CT + TT (Rec)	39(78.00%)	30(76.92%)		1.16(0.49-2.75)	0.737		1.59(0.59-4.34)	0.361		1.08(0.39-2.97)	0.883
rs4779584	CC	34(68.00%)	25(64.10%)		1.00			1.00			1.00	
T	CT	15(30.00%)	14(35.90%)		0.86(0.40-1.83)	0.692		0.98(0.44-2.17)	0.954		1.21(0.49-3.00)	0.674
	TT	1(2.00%)	0(0.00%)	0.38	0.41(0.04-3.71)	0.426	0.63	-	-	0.75	-	-
	CT + TT, vs. CC (Dom)	16(32.00%)	14(35.90%)		0.80(0.38-1.68)	0.558		0.85(0.39-1.87)	0.685		1.15(0.47-2.82)	0.760
	TT, vs. CT + CC (Rec)	49(98.00%)	39(100.0%)		0.43(0.98-2.36)	0.451		-	-		-	-
rs4939827	CC	23(46.00%)	5(12.82%)		1.00			1.00			1.00	
T	CT	16(32.00%)	19(48.72%)		0.17(0.07-0.40)	0.0001		1.05(0.33-3.32)	0.928		5.61(1.70-18.40)	0.004
	TT	11(22.00%)	15(38.46%)	0.01	0.28(0.10-0.76)	0.012	0.15	2.01(0.60-6.81)	0.258	0.001	6.16(1.75-21.68)	0.005
	CT + TT, vs. CC (Dom)	27(54.00%)	34(87.18%)		0.20(0.09-0.45)	0.0001		1.33(0.44-3.99)	0.608		5.63(1.93-17.60)	0.002
	TT, vs. CT + CC (Rec)	39(78.00%)	24(61.54%)		0.84(0.36-1.92)	0.676		1.94(0.86-4.34)	0.109		2.15(0.83-5.53)	0.114
rs6983267	TT	16(32.00%)	4(10.53%)		1.00			1.00			1.00	
G	GT	19(38.00%)	19(50.00%)		0.48(0.20-1.15)	0.101		1.81(0.54-6.04)	0.336		3.97(1.09-14.40)	0.036
	GG	15(30.00%)	15(39.47%)	0.39	0.66(0.26-1.68)	0.379	0.14	2.57(0.72-9.12)	0.144	0.07	4.42(1.14-17.05)	0.031
	GT + GG, vs. TT (Dom)	34(68.00%)	34(89.47%)		0.54(0.25-1.21)	0.136		2.07(0.65-6.59)	0.218		4.15(1.22-14.08)	0.022
	GG, vs. GT + TT (Rec)	35(70.00%)	23(60.53%)		1.04(0.48-2.25)	0.911		1.63(0.72-3.67)	0.243		1.67(0.67-4.20)	0.274
rs9929218	GG	22(44.00%)	21(53.85%)		1.00			1.00			1.00	
A	AG	24(48.00%)	14(35.90%)		0.82(0.39-1.70)	0.588		0.50(0.18-2.28)	0.490		0.60(0.24-1.48)	0.268
	AA	4(8.00%)	4(10.26%)	0.38	0.64(0.18-2.30)	0.496	0.13	0.64(0.18-2.28)	0.490	0.48	1.07(0.23-4.90)	0.927
	AG + AA, vs. GG (Dom)	28(56.00%)	18.15(%)		0.78(0.38-1.59)	0.502		0.53(0.25-1.13)	0.098		0.66(0.28-1.57)	0.352
	AA, vs. AG + GG (Rec)	46(92.00%)	35(89.74%)		0.72(0.21-2.42)	0.591		0.89(0.26-3.02)	0.857		1.36(0.31-5.87)	0.680

* OR adjusted for gender and age groups, ** OR adjusted for gender and age at diagnosis

in a pilot group of 92 CRC cases and 96 healthy controls.

To our best knowledge, this is the first study aimed to test these associations in a Romanian sample. The primary aim was to determine whether any of these polymorphisms associate with the CRC risk and to observe their effects (if any) in our population. Secondary, we investigate the possible association with several phenotypes of the disease in our group.

SNPs rs10795668, rs16892766 were reported by Thomlinson and collaborators in a follow up study with three stages of replication (20). For rs16892766 the homozyous state was associated with an increased risk of CRC and the effect was significantly stronger in younger patients. For rs10795668 the susceptibility allele was more prevalent in rectal tumors indicating a site specificity similar to those found in our analysis. The role of these two SNPs is controversial: while rs16892766 lies in a linkage disequilibrium (LD) block that includes a gene which regulates cell growth and viability (EIF3H), the function of rs10795668 is more difficult to elucidate since the SNP is located in a region without proven protein-coding transcripts (21).

In another phased design GWAS Tenesa et al. identified a new risk locus on 11q23 tagged by SNP rs3802842 located close to a gene called POU2AF1 and replicated the modest but highly significant associations between rs4939827 and CRC risk reported by Broderick (22,23). Substantial population-specific differences were reported with different allelic effects. Furthermore, Tenesa et al found notable site-specific differences in risk, namely greater risk for rectal than colon cancer, associated with both SNPs rs3802842 (11q23 locus) and rs4939827 (18q21) when comparing colon and rectal cancers (23). At the same time, Curtin et al. (24) reported findings in the opposite direction for rs4939827, while Broderick et al. (22) observed no difference by site in the London cohort.

Our results concerning SNP rs4939827 are in the same direction to those found by Tenesa, while for SNP rs3802842 our findings are opposite those indicated by Tenesa. In spite of these discrepancies, largely due to our modest sample size, additional site specific investigations are needed.

The rs4939827 maps to SMAD7, an intercellular antagonist of transforming growth factor TGF β1 signaling, which binds to the receptor complex and is able to block the activation of downstream signaling events (25). The TGF β1 signaling pathway function as both a tumor suppressor in early stage cancers and as an oncogene as well in advanced cases and metastasis.

The recent meta-analysis published by Houlston et al. (16) revealed four new CRC loci. The second strongest association was for rs4444235 on 14q22.2. It maps 9.4 kb from the transcription start site of the gene BMP4 which belongs to (TGF) β family and therefore could play an important role in CRC (26). Our results points to the same increased risk for rs4444235 confirmed in all analysis stratified by tumor location.

The increased risk of CRC for rs4779584 was initially found in CRC cases selected for family history. The SNP location close to known SCG5 and GREM1 genes could explain a role in the tumorigenic process (27).

Although the inverse association for rs4779584 found in our case-control series cannot be explained, our results presented for rs9929218 are in the same direction to those reported by Houlston (16).

The 8q24 region was the first susceptibility risk locus identified to be associated with prostate and colorectal cancer. The region is tagged by rs6983267 and rs10505477 and is in the close vicinity to the oncogene MYC that is known for its role in colon cancer biology. Moreover, a separate analysis in additional cancer-free adenoma provided evidence that rs6983267 is associated with an elevated risk of adenoma development (28). Pooled analysis of genetic variation at 8q24 and CRC risk potentially indicate a role in accelerating adenoma formation (29). The development of adenoma is known to be more prevalent in the cecum and rectum compared to the rest of the large bowl (30) Thus, it is possible that the strong association with rectal, rather than colon cancer, observed in our study for this SNP links to the neoplastic transformation of adenoma.

Strengths of this study include that it was conducted in a highly homogenous population of a single ethnic group, with a design allowing accurate characterization of disease phenotype. Furthermore, stringent genotyping quality control was applied as a check on genotyping accuracy.

The major limitation of our study was the power to detect association between SNPs and CRC risk given the small sample size. Thus, inevitably we may have missed genotype-phenotype associations and our observations do not allow enough statistically significant conclusions.

Conclusions

This is the first SNP replication study of CRC risk association in a Romanian sample.

Our study confirms previously-identified associations with colon and rectal cancer risk for rs6983267, rs4939827, rs3802842, rs4444235, rs10795668 (five out of eight SNPs investigated) and underlines the necessity of extensive replication using larger samples.

Acknowledgements

This work was supported by CNCSIS–UEFISCSU, project number PNII – IDEI code: ID-1184/2008, and by deCODE Genetics, Reykjavik, Iceland (Collaboration Agreement, April 9, 2008).

We gratefully acknowledge the contribution of Thorunn Rafnar for genotyping data acquisition.

References

1. World Health Organization, World Cancer Report (eds. Stewart B.W. & Kleihues P.), 13, IARC, Lyon, 2003.
2. Ferlay J, Autier P, Boniol M, Heanue M, Colombet P.

Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol. 2007;18:581-92.
3. European Cancer Observatory. Cancer Fact Sheet. IARC 2006 (http;//eu-cancer.iarc.fr/2-cancer-fact-sheets.html.en).
4. Valean S, Armean P, Resteman S, Nagy G, Muresan A, Mircea PA. Cancer mortality in Romania, 1955-2004. Digestive sites: esophagus, stomach, colon and rectum, pancreas, liver, gallbladder and biliary tree. J Gastrointestin Liver Dis. 2008;17(1):9-14.
5. Boyle P, Ferlay J. Mortality and survival in breast and colorectal cancer. Nat Clin Pract Oncol. 2005;2:424-5.
6. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004;10(8):789-99.
7. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer-analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000;343(2):78-85.
8. Aaltonen L, Johns L, Järvinen H, Mecklin JP, Houlston R. Explaining the familial colorectal cancer risk associated with mismatch repair (MMR)-deficient and MMR-stable tumors. Clin Cancer Res. 2007;13(1):356-61.
9. Gruber SB, Ellis NA, Rennert G, Offit K. BLM heterozygosity and the risk of colorectal cancer. Science. 2002;297(5589):2013.
10. Laken SJ, Petersen GM, Gruber SB, et al. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. Nat Genet. 1997;17(1):79-83.
11. Bost B, de Vienne D, Hospital F, Moreau L, Dillmann C. Genetic and nongenetic bases for the L- shaped of quantitative trait loci effects. Genetics. 2001;157:1773-87.
12. Chung CC, Magalhaes WCS, Gonzales-Bosquet G, Chanock SJ. Genome wide association studies in cancer – current and future directions. Carcinogenesis. 2010;31 (1):111-20.
13. Sebastiani P, Lazarus R, Weiss ST, Kunkel LM, Kohane IS, Ramoni MF. Minimal haplotype tagging. Proc Natl Acad Sci USA.2003;100:9900-5.
14. Gibbs RA, Belmont JW, Hardenbol P et al. The International HapMap project. Nature.2003;426:796-8.
15. Tenesa A, Malcolm GD. New insights into the aetiology of colorectal cancer from genome-wide association studies. Nature Reviews Genetics.2009;10:353-8.
16. 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(12):1426-35.
17. International Union Against Cancer (UICC). TNM Classification of Malignant Tumours, Sixth Edition. Wiley-Liss (New-York), 2002, 72-76.
18. International Classification of Diseases for Oncology, Third Edition, WHO Geneva, 2000, 49.
19. Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TA. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies. Accessed 20.09.2010.
20. Tomlinson IPM, 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;40:623-30.
21. Picelli S, von Holst S, Wessendorf P. The continuing search for predisposing colorectal cancer variants. Cancer Genomics Proteomics. 2009;6:305-16.
22. 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;39:1315-7.
23. 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;40:631-7.
24. Curtin K, Lin WI, George R et al. Meta association of colorectal cancers confirms risk alleles at 8q24 and 18q21. Cancer Epidemiol Biomarkers Prev. 2009;18:616-21.
25. ten Dijke P, Hill CS. New insights into TGF-beta-Smad signalling. Trends Biochem Sci. 2004;29:265-73.
26. Kim JS, Crooks H, Dracheva T et al. Oncogenic beta-catenin is required for bone morphogenetic protein 4 expression in humn cancer cells. Cancer Res. 2002;62:2744-8.
27. Jaeger E, Webb E, Howarth K at al. Common genetic variants at CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. Nat Genet. 2008;40: 26-8.
28. 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;39:984-8.
29. Berndt S, Potter JD, Hazra A, et al. Pooled analysis of genetic variation at chromosome 8q24 and colorectal neoplasia risk. Hum Mol Genet. 2008;17:2665-72.
30. Makinen MJ, George SMC, Jernvall P, Makela J, Vihko P, Karttunen TJ. Colorectal carcinoma associated with serrated adenoma – prevalence, histological features, and prognosis. J Pathol. 2001;193(3):286-94.