Mutation Research 648 (2008) 40–45

Contents lists available at ScienceDirect

Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis

journal homepage: www.elsevier.com/locate/molmut
Community address: www.elsevier.com/locate/mutres

Do polymorphisms and haplotypes of mismatch repair genes modulate risk of sporadic colorectal cancer?

E. Tulupova[a,b], R. Kumar[c], M. Hanova[a,b], J. Slyskova[a,b], B. Pardini[a], V. Polakova[a], A. Naccarati[a], L. Vodickova[a,d], J. Novotny[e], J. Halamkova[f], K. Hemminki[c,g], P. Vodicka[a,*]

[a] *Institute of Experimental Medicine, Academy of Sciences of the Czech Republic, Videnska 1083, 14200 Prague 4, Czech Republic*
[b] *Faculty of Science, Charles University, Prague, Czech Republic*
[c] *German Cancer Research Center (DKFZ), Heidelberg, Germany*
[d] *National Institute of Public Health, Prague, Czech Republic*
[e] *1st Medical Faculty, Charles University, Prague, Czech Republic*
[f] *Faculty Hospital Brno, Brno, Czech Republic*
[g] *Center for Family Medicine, Karolinska Institute, Huddinge, Sweden*

ARTICLE INFO

Article history:
Received 11 June 2008
Received in revised form 12 August 2008
Accepted 11 September 2008
Available online 21 September 2008

Keywords:
DNA mismatch repair
Genetic polymorphism
Haplotype analysis
Colorectal cancer susceptibility

ABSTRACT

The Czech Republic presents one of the highest incidences of colorectal cancer in the world. We genotyped 10 single nucleotide polymorphisms in five DNA mismatch repair genes in 614 colorectal cancer cases and 614 matched controls from this country. The carriers of T-allele of the *hMSH6*-556G > T polymorphism were at increased risk of colorectal cancer (OR 1.29; 95% CI 1.02–1.62). The stratification of data showed that risk associated with the polymorphism was confined to rectal cancer (OR 1.42; 95% CI 1.03–1.95). The A-allele of the Ex1 – 145G > A polymorphism in the *hMSH6* gene was associated with a decreased risk of colorectal cancer (OR 0.76; 95% CI 0.60–0.98). The C-allele of the IVS4-101G > C polymorphism in *hMSH6* was associated with an increased risk of colon cancer (OR 1.34; 95% CI 1.03–1.74). The carriers of the variant allele for the polymorphism IVS9-1406C > T in *hMLH1* exhibited a decreased risk of rectal cancer (OR 0.71; 95% CI 0.51–0.98). We observed a differential distribution of haplotypes based on three *hMSH6* polymorphisms (-556G > T-Ex1 – 145G > A-IVS4-101G > C) in the cases and controls (global $P = 0.02$). The TAG haplotype was associated with a decreased risk of colorectal cancer (OR 0.74; 95% CI 0.59–0.92), whereas the most frequent haplotype GGG was associated with increased risk of rectal cancer (OR 1.32; 95% CI 1.05–1.65). However, multiple hypotheses testing diminishes a statistical significance of above associations. Our data suggest a limited role for the investigated individual variants in mismatch repair genes for the susceptibility to the disease. The haplotypes covering *hMSH6* gene may, however, be involved in risk modulation in this population.

1. Introduction

The mismatch repair (MMR) system plays a key role in maintenance of genomic stability through removal of mismatched nucleotide pairs and insertion/deletion heterologies generated during DNA replication. The components of DNA MMR machinery are also involved in cell cycle arrest and induction of apoptosis in response to different types of DNA damage [1,2]. Inactivation of MMR is associated with hereditary and sporadic human cancers either through occurrence of mutations or through epigenetic silencing of the component genes, leading to microsatellite instability (MSI) [3]. Both germline mutations and hypermethylation in MMR genes have been reported in familial/hereditary forms of colorectal cancer (CRC). In the majority of sporadic CRC with MSI, a hypermethylation of the *hMLH1* promoter is considered to be the main mechanism of MMR inactivation. However, the impact of genetic and epigenetic mutations of MMR genes for the risk of sporadic form of CRC is currently quite indeterminate [4]. It is possible that the common variants in relevant genes encoding DNA MMR enzymes also impact the risk of the sporadic form of the disease on a population level.

The recent discovery, through genome-wide association scans, of a number of polymorphisms and loci associated with the disease susceptibility has provided insight into the role of low penetrance variants in the disease aetiology [5,6]. The latter report provides further evidence for "common disease—common variant" model of CRC predisposition.

* Corresponding author. Tel.: +420 241 062 694; fax: +420 241 062 782.
E-mail address: pvodicka@biomed.cas.cz (P. Vodicka).

0027-5107/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrfmmm.2008.09.005

The genetic phenomena associated with MMR insufficiency in CRC, such as MSI and loss of heterozygosity leading to "mutator phenotype", have been extensively studied [2,7]. The investigations on the role of genetic polymorphisms for the disease susceptibility in population have lagged for a number of reasons that include lack of proper design with sufficient sample sizes. Moreover, the functional relevance of majority of polymorphisms in the genes involved in MMR is not known. Though recent studies suggest an influence of single nucleotide polymorphisms (SNPs) on biochemical interaction between components of the MMR pathway or on epigenetic mediated functional regulation [8,9]. In view of the well-defined role of defective MMR in CRC, it is important to determine if common variants in the involved genes affect susceptibility to the disease at a population level.

CRC is one of the most common cancers in developed countries [10], accounting annually for 1,200,000 newly diagnosed cases and over 525,000 deaths worldwide [11]. In a hospital-based case-control study on the Czech population we investigated 10 SNPs in MMR genes *hMLH1*, *hMSH2*, *hMSH3*, *hMSH6*, and *hEXO1*. The selected variants included tagging SNPs for the *hMLH1* and *hMSH6* genes. CRC represent a major health problem in the Czech Republic with incidence rate being one of the highest in the world [11].

2. Material and methods

2.1. Study population

The study population consisted of 614 CRC patients and 614 hospital-based controls, matched for sex and age. The eligibility criteria for participation in the study were that subjects were aged 29 years or more, were of Czech origin, and consented to provide biological samples for genetic analysis. The cases were recruited from nine oncological departments (two in Prague, one in Benesov, Brno, Liberec, Ples, Pribram, Usti nad Labem, and Zlin), covering nearly the entire country, between September 2004 and February 2006. The final inclusion of cases into the study was based on the histological confirmation of the diagnosis. During the study period, a total of 968 cases with CRC provided blood samples in above hospitals. Sixteen individuals were excluded as they fulfilled the Amsterdam criteria I and II for hereditary CRC. The data of 952 subjects were used for matching with controls.

Controls were recruited in parallel among individuals admitted to five large gastroenterological departments all over the Czech Republic (Prague, Brno, Jihlava, Liberec, and Pribram). Each control subject was undergoing colonoscopy for various gastrointestinal complaints. The reasons for colonoscopical investigation were (i) macroscopic bleeding; (ii) positive Fecal Occult Blood Test (FOBT); (iii) abdominal pain of unknown origin. Due to the high incidence of CRC in the Czech Republic, colonoscopy is largely recommended and practiced. The control group was composed of subjects with negative colonoscopic results for malignancy or idiopathic bowel diseases [12]. To reduce selection bias, we included as controls only those subjects with no previous diagnosis and without any chronic disease or those repeatedly admitted to hospital and modified habits because of their disease. Among 739 recruited controls, a total of 663 (89.7%) were used for matching for sex and age with CRC patients. As a result, 614 case-control pairs were formed. Thus, 338 cases and 49 controls not fitting into the pairs or with incomplete lifestyle and potential risk factor information were excluded from initial groups. The sex distribution in excluded controls was similar to those included.

The participating subjects were properly informed and provided a written consent and approval for genetic analysis. The study-design was approved by the Ethical Committee of the Institute of Experimental Medicine, Prague, Czech Republic.

2.2. Interviews

Cases and controls were interviewed by trained personnel using a structured questionnaire to determine demographic characteristics and potential risk factors for CRC. Study subjects provided information on their lifestyle habits, body mass index (BMI), diabetes, and family/personal history of cancer. Lifelong or long-term (at least six consecutive months) drug use was included in the questionnaire.

2.3. SNP selection and genotyping

For this study we chose 10 SNPs in MMR genes *hMLH1*, *hMSH2*, *hMSH3*, *hMSH6* and *hEXO1.* The selected polymorphisms were located in coding and non-coding regions of the genes and exerted a possible functional effect according to results of association and/or *in vitro* studies. The polymorphism studied were: two SNPs in non-coding regions of the *hMLH1* gene (the -93G > A in promoter (rs1800734) and the IVS9-1406C > T (rs4647269) polymorphism in intron 9); two nucleotide changes in the *hMSH2* (the SNP located in the end of intron 12 IVS12-6T > C (rs2303428) and a missense variant Ex6 + 23G > A (rs4987188) coding the Gly322Asp amino acid change); two exonic SNPs in the *hMSH3* gene (Ex4-100G > A (rs1805355, Pro231Pro) in exon 4 and Ex23 + 3G > A (rs26279) in exon 23 coding the Ala1045Thr amino acid substitution); two SNPs in the *hMSH6* gene -556G > T (rs3136228) located in the promoter region and Ex1 − 145G > A (rs1042821, Gly39Glu) in exon 1 and a SNP in the *EXO1* exon 12 (Ex12 + 49C > T; rs4149963) coding Thr439Met amino acid substitution. Two tagging SNPs the IVS9-1406C > T (rs4647269) polymorphism in the *hMLH1* gene and the IVS4-101G > C (rs2072447) in the *hMSH6* gene, capturing 21 and 2 SNPs, respectively, were selected based on Hapmap data (www.hapmap.org) for corresponding haplotypes reconstruction. All studied SNPs (with exception of more rare Ex6 + 23G > A polymorphism in *hMSH2*, MAF = 0.03) are common in Caucasians (MAF 0.06–0.53) according to Hapmap and SNP500 databases.

DNA was isolated from peripheral blood lymphocytes of CRC patients and control subjects using phenol–chloroform procedure and stored at −20 °C. The purity was assessed by spectrophotometry. Pre-designed assays for allelic discrimination (Taqman) were used to genotype eight out of the ten selected SNPs. The amplification for genotyping was performed on an AB 7500 Real-Time PCR (Applied Biosystems, Foster City, USA) system. For polymorphisms Ex6 + 23G > A (rs4987188) in the *hMSH2* gene and Ex1 − 145G > A (rs1042821) in the *hMSH6* gene we used a PCR-RFLP method due to non-availability of pre-designed assays for allelic discrimination. The genotyping for the Ex6 + 23G > A polymorphism in the *hMSH2* gene was carried out as previously described [13]. For genotyping the Ex1 − 145 G > A polymorphism in the *hMSH6* gene following primers were designed using Primer3 v.0.4.0 (http://frodo.wi.mit.edu): sense 5′-AGATGCGGTGCTTTTAGGAG-3′, antisense 5′-CCCTCCGTTGAGGTTCTTC-3′. Each PCR reaction (20 μl) contained 50 ng of genomic DNA, 1 U of HOT FIREPol DNA polymerase I, 25 mM Tris–HCl, 50 mM KCl, 2.5 mM $MgCl_2$, 300 μM dNTPs, and 0.3 μM of each primer. Thermal cycle conditions were as suggested by manufacturer (Odex, Lucca, Italy): 95 °C 15 min, 20 cycles (95 °C 1 min, 68 °C 30 s—1 °C/cycle, 72 °C 1.5 min), 20 cycles (95 °C 1 min, 51 °C 30 s, 72 °C 1.5 min), 72 °C 10 min. The 289 bp PCR product was digested with 6 U of restriction enzyme SmaI (New England BioLabs Inc., USA) overnight. The G allele was digested into 114 and 175 bp fragments, whereas the A variant remained intact. The digested PCR products were resolved on 2% agarose gel with ethidium bromide and visualized under UV light.

The genotype screening was performed simultaneously for cases and controls. Approximately 10% of the samples were re-genotyped for each of ten polymorphisms analysed to confirm the initial results.

2.4. Statistical analyses

Genotype distribution for each polymorphism in controls was tested for deviation from the Hardy–Weinberg equilibrium by Pearson chi-square test. Differences in baseline socio-demographic characteristics between cases and controls were analysed using chi-square test and Student's *t*-test. Multivariate logistic regressions were used to examine the association between each genotype and risk of selected health endpoints (all CRC, colon cancer, and rectal cancer separately). Odds ratio (OR) and 95% confidence interval (CI) were estimated including gender and age as covariates. Separate analyses were carried out following the stratifications for smoking and for tumor localization.

The haplotype frequencies in cases and controls were inferred with the SAS/Genetics software module. The analysis was carried out to examine the phase of *hMLH1*, *hMSH2*, *hMSH3* and *hMSH6* polymorphisms, using the expectation–maximization algorithm to generate maximum likelihood estimates of haplotype frequencies.

Linkage disequilibrium was calculated with Haploview software (www.broad.mit.edu/mpg/haploview/documentation.php).

3. Results

3.1. Study population

A modulating role of 10 polymorphisms in MMR genes and their haplotypes on a risk of CRC (as well as colon and rectal cancers separately) was investigated using a case-control approach (Table 1). Sex and age were chosen as main criteria for matching cases and controls. A preliminary analysis revealed that there were no major significant differences for the considered covariates between the patients and controls, with the exception of a moderately different distribution of ex-smokers (defined as individuals who quit smoking less than 5 years prior to the sampling). A statistically significant difference was observed in dietary habits between the cases and controls with prevailing vegetarians among cases. Moreover, controls showed a higher percentage of individuals with higher degree of education as compared to cases.

Table 1
Characteristics of CRC patients and control subjects.

	Cases ($N = 614$)	Controls ($N = 614$)
Gender		
Males	343	343
Females	271	271
Age at diagnosis (years)		
Mean ± S.D.	58.6 ± 10.4	57.8 ± 12.4
Median	58	58
Range	26–84	29–85
Diagnosis		
Colon cancer	217	–
Sigmoidal	156	–
Rectal cancer	241	–
Colon + Sigm.	373	–
Cancer family history (%)		
Positive history	54.2	59.7
Negative history	45.8	40.3
Smoking status (%)		
Non-smokers	51.1	53.4
Ex-smokers >5 years	22.1	20.9
Ex-smokers <5 years	10.6	4.4
Current smokers	16.2	21.3
No of cigarettes (%)		
<20 cig/day	58.0	59.4
>20 cig/day	42.0	40.6
Alcohol (%)		
No	46.5	42.6
Yes	53.5	57.4
Diet (%)		
Strictly vegeterian	12.5	6.6
Otherwise	87.5	93.4
Living (%)		
City	56.7	54.6
City + Country	14.7	19.5
Country	28.6	25.9
Education (%)		
Basic school	34.8	28.5
High school	51.5	53.2
University	13.7	18.3
Body mass index		
Mean ± S.D.	26.7 ± 4.4	27.0 ± 4.5
<18.5 (%)	1.5	0.4
18.5–24.9 (%)	36.6	35.0
25.0–29.9 (%)	43.1	43.1
30.0–39.9 (%)	17.8	19.9
>40 (%)	1.0	1.6

3.2. Genotype analyses

The genotype distribution of polymorphisms analysed in this study did not show statistically significant deviation from the Hardy–Weinberg equilibrium. The data on allele and genotype distributions in CRC cases and controls, and the sex and age adjusted ORs for associated risk are presented in Table 2. The minor allelic frequencies (MAF) for polymorphisms in the MMR genes in the Czech population were in concordance with those reported for Caucasian population in the public HAPMAP (http://www.hapmap.org/citinghapmap.html.en) or SNP500 databases (http://snp500cancer.nci.nih.gov/terms_ethnic.cmf).

Two polymorphisms in the *hMSH6* gene showed an association with altered risk of CRC ($P \leq 0.05$). Carriers of variant allele for the -556G > T polymorphism in the promoter region of the gene were at an increased risk of CRC (OR 1.29; 95% CI 1.02–1.62; $P = 0.04$), whereas the carriers of the variant A-allele for the -145G > A polymorphism in the *hMSH6* exon 1 were at a decreased risk of CRC.

Table 2
Distribution of MMR genotypes in CRC patients and controls and results of unconditional logistic regression analysis[a].

Genotype	Cases[b]	Controls[b]	OR	CI 95%	P[c]
hMLH1-93G > A (rs1800734)					
GG	359	365	1 (ref)	–	0.75
GA	216	209	1.10	0.86–1.40	
AA	34	37	1.00	0.61–1.64	
GA + AA	250	246	1.10	0.86–1.36	0.50
hMLH1 IVS9-1406C > T (rs4647269)					
CC	193	172	1 (ref)	–	0.21
CT	304	317	0.83	0.64–1.08	
TT	107	119	0.75	0.54–1.05	
CT + TT	411	436	0.81	0.63–1.04	0.10
hMSH2 IVS12-6T > C (rs2303428)					
TT	505	516	1 (ref)	–	0.59
TC	102	89	1.16	0.85–1.58	
CC	4	3	1.46	0.32–6.57	
TC + CC	106	92	1.17	0.86–1.59	0.32
hMSH2 Ex6 + 23G > A (rs4987188)					
GG	586	588	1(ref)	–	0.70
GA	20	22	0.89	0.47–1.66	
AA	0	0	–	–	–
hMSH3 Ex4-100G > A (rs1805355)					
GG	507	531	1 (ref)	–	0.56
GA	92	73	1.36	0.97–1.91	
AA	7	6	1.28	0.43–3.84	
GA + AA	99	79	1.36	0.98–1.88	0.07
hMSH3 Ex23 + 3G > A (rs26279)					
GG	302	307	1 (ref)	–	0.74
GA	265	256	1.07	0.85–1.36	
AA	41	47	0.93	0.59–1.46	
GA + AA	306	303	1.05	0.84–1.32	0.67
hMSH6-556G > T (rs3136228)					
GG	225	254	1 (ref)	–	0.10
GT	293	272	1.26	0.98–1.62	
TT	94	85	1.36	0.96–1.93	
GT + TT	387	357	**1.29**	**1.02–1.62**	**0.04**
hMSH6 Ex1 − 145G > A (rs1042821)					
GG	428	399	1(ref)	–	0.08
GA	153	181	0.78	0.61–1.01	
AA	20	29	0.63	0.35–1.14	
GA + AA	173	210	**0.76**	**0.60–0.98**	**0.03**
hMSH6 IVS4-101G > C (rs2072447)					
GG	278	306	1 (ref)	–	0.29
GC	268	253	1.18	0.93–1.50	
CC	59	51	1.26	0.83–1.90	
GC + CC	327	304	1.20	0.95–1.50	0.13
hEXO1 Ex12 + 49C > T (rs4149963)					
CC	479	497	1 (ref)	–	0.43
CT	123	106	1.21	0.90–1.62	
TT	7	7	0.91	0.30–2.73	
CT + TT	130	113	1.19	0.90–1.58	0.23

OR, odds ratio; CI 95%, confidence interval; ref., reference.

[a] Adjusted for sex and age.

[b] Numbers may not add up to 100% of subjects due to genotyping failure. All samples that did not give a reliable result in the first round of genotyping were resubmitted to up to three additional rounds of genotyping. Data points that were still not filled after this procedure were left blank.

[c] Bold values indicate $P \leq 0.05$; no significant difference in genotype distributions were recorded after correction for multiple hypotheses testing.

The data stratification based on the tumor location showed that the increased risk associated with the variant allele of the -556G > T polymorphism in the *hMSH6* promoter was more pronounced for rectal cancer (OR 1.42; 95% CI 1.03–1.95; $P = 0.03$) than for colon cancer (OR 1.21; 95% CI 0.92–1.58; $P = 0.17$; Table 3). In addition, we found that in the *hMSH6* gene the variant allele of the intronic IVS4-101G > C SNP was associated with increased risk of colon cancer (OR

Table 3
Distribution of genotypes in MMR genes and results of unconditional logistic regression analysis[a] for CRC patients stratified for tumor location and controls.

Genotype	Controls[b]	Colon[b]	OR	CI 95%	*P*	Rectum[b]	OR	CI 95%	*P*[c]
hMLH1-93G > A									
GG	365	142				131			
GA + AA	246	228	0.96	0.73–1.25	0.74	108	1.31	0.96–1.78	0.09
hMLH1 IVS4-101G > C									
CC	172	111				82			
CT + TT	436	257	0.89	0.67–1.19	0.43	154	**0.71**	**0.51–0.98**	**0.04**
hMSH2 IVS12-6T > C									
TT	516	309				196			
TC + CC	92	62	1.12	0.79–1.6	0.53	44	1.24	0.83–1.85	0.30
hMSH2 Ex6 + 23G > A									
AA	588	356				230			
GA	22	13	0.92	0.45–1.89	0.82	7	0.82	0.35–1.95	0.66
hMSH3 Ex4-100G > A									
GG	531	309				198			
GA + AA	79	59	1.33	0.92–1.93	0.13	40	1.38	0.9–2.11	0.14
hMSH3 Ex23 + 3G > A									
GG	307	196				106			
GA + AA	303	172	0.90	0.69–1.17	0.44	134	1.34	0.98–1.81	0.06
hMSH6-556G > T									
GG	254	144				81			
GT + TT	357	228	1.21	0.92–1.58	0.17	159	**1.42**	**1.03–1.95**	**0.03**
hMSH6 Ex1 − 145G > A									
GG	399	259				169			
GA + AA	210	107	0.78	0.59–1.03	0.08	66	0.74	0.53–1.03	0.08
hMSH6 IVS4-101G > C									
GG	306	158				120			
GC + CC	304	209	**1.34**	**1.03–1.74**	**0.03**	118	1.02	0.75–1.38	0.92
hEXO1 Ex12 + 49C > T									
CC	497	285				194			
CT + TT	113	85	1.31	0.95–1.81	0.10	45	1.03	0.7–1.51	0.89

OR, odds ratio; CI, 95%, confidence interval; ref., reference.

[a] Adjusted for sex and age.

[b] Numbers may not add up to 100% of subjects due to genotyping failure. All samples that did not give a reliable result in the first round of genotyping were resubmitted to up to three additional rounds of genotyping. Data points that were still not filled after this procedure were left blank.

[c] Bold values indicate $P \leq 0.05$; no significant difference in genotype distributions were recorded after correction for multiple hypotheses testing.

1.34; 95% CI 1.03–1.74; $P = 0.03$). On the other hand, carriers of the T-allele for the IVS9-1406C > T polymorphic variant *hMLH1* exhibited a decreased risk of rectal cancer (OR 0.71; 95% CI 0.51–0.98; $P = 0.04$). However, by correcting for multiple hypotheses testing the above associations are not convincingly significant.

The data analysis did not show any association with either smoking habit or family history of CRC with any of the polymorphisms in the MMR genes included in this study (data not shown).

3.3. Haplotype analyses

By constructing haplotypes for four MMR genes, we observed a statistically significant difference in the *hMSH6* haplotype distribution between the CRC cases and controls (global test $P = 0.02$). The TAG haplotype (-556G > T-Ex1 − 145G > A-IVS4-101G > C) was associated with decreased risk of CRC in the studied population (Table 4; OR 0.74; 95% CI 0.59–0.92). None of the haplotypes in the *hMLH1*, *hMSH2*, or *hMSH3* genes was significantly associated with risk of CRC (global tests for *hMLH1* $P = 0.11$, *hMSH2* $P = 0.59$, and for *hMSH3* $P = 0.31$).

The analysis of *hMSH6* haplotype distribution after the stratification for tumor localization (Table 5) revealed significant differences between controls and patients with rectal cancer in particular (global *P*-value 0.03). Global *P*-value for haplotype effect on colon cancer (0.12) did not reflect significant differences in haplotype distribution between the cases and controls. The TAG haplotype for *hMSH6* (-556G > T-Ex1 − 145G > A-IVS4-101G > C) was associated with significantly decreased risk in both colon and rectal cancer patients (Table 5; OR 0.74; 95% CI 0.51–0.96 and OR 0.73; 95% CI 0.53–1.00, respectively). The GGG haplotype was found to be exclusively associated with increased risk of rectal cancer (OR 1.32; 95% CI 1.05–1.65). None of the MMR haplotypes was significantly associated with the familial aggregation of the disease.

The linkage disequilibrium (LD) between the three loci analysed in the *hMSH6* gene (-556G > T-Ex1 − 145G > A-IVS4-101G > C) is shown in Fig. 1.

Table 4
Distribution of *hMSH6* haplotypes in patients with CRC and in controls.

Haplotype[a]	Cases N[b]	Controls N[b]	OR	CI 95%	*P*-value[c]
TGG	180	218	0.84	0.68–1.05	**0.02**
TGC	350	322	1.16	0.96–1.39	
TAG	165	217	**0.74**	**0.59–0.92**	
TAC	17	16	1.07	0.51–2.25	
GGG	452	413	1.16	0.98–1.30	
GGC	12	13	0.93	0.4–2.18	
GAG	8	3	2.70	0.65–12.86	

[a] *hMSH6*-556G > T-Ex1 − 145G > A-IVS4-101G > C.

[b] *N* is the number of alleles. Because each individual has two alleles, the total number of alleles will be twice the total number of individuals. Individuals missing haplotype data were not included in the analyses.

[c] Global *P*-value for haplotype effect calculated from χ^2 test. Significant ORs are in bold.

Table 5
Distribution of *hMSH6* haplotypes in CRC patients stratified for tumor location and in controls.

Haplotypes[a]	Controls *N*[b]	Colon cancer cases *N*[b]	OR (CI 95%)[c]	*P*-value[c]	Rectum cancer cases *N*[b]	OR (CI 95%)[c]	*P*-value[c]
TGG	218	114	1.03 (0.80–1.33)	0.12	66	1.35 (0.97–1.85)	**0.03**
TGC	322	217	1.17 (0.95–1.45)		132	1.06 (0.83–1.36)	
TAG	217	101	**0.74 (0.51–0.96)**		64	**0.73 (0.53–1.00)**	
TAC	16	13	1.36 (0.61–3.00)		4	0.65 (0.18–2.08)	
GGG	413	264	1.10 (0.90–1.34)		188	**1.32 (1.05–1.65)**	
GGC	13	9	1.15 (0.45–2.90)		3	0.60 (0.14–2.25)	
GAG	3	4	2.23 (0.42–12.52)		4	3.50 (0.66–19.69)	

[a] *hMSH6*-556G > T-Ex1 – 145G > A-IVS4-101G > C.
[b] *N* is the number of alleles. Because each individual has two alleles, the total number of alleles will be twice the total number of individuals. Individuals with missing haplotype data were not included in the analyses.
[c] Global *P*-value for haplotype effect calculated from χ^2 test. Significant ORs are in bold.

-556G>T	Exl-145G>A	IVS4-101G>C	
1	0.92 (0.11)	0.89 (0.22)	-556G>T
	1	0.65 (0.04)	Exl-145G>A
		1	IVS4-101G>C

Fig. 1. Linkage disequilibrium $|D'|$ (R^2) between polymorphisms in the *hMSH6* gene.

4. Discussion

We investigated variants in the genes encoding proteins involved in DNA MMR and their role in modulating susceptibility to CRC using a case-control approach. An analysis of questionnaire-based lifestyle data revealed a significant difference in dietary habits in cases with prevailing vegetarians as compared to controls. However, hardly any far-reaching conclusions may be drawn based on this observation due to the subjectivity of questionnaires which decreases reliability of the given information. Moreover, as discussed recently [14], questionnaire-based observations introduce profound bias due to dietary recall and influence of confounders. Dietary habits during lifetime may further be modulated, as consequence of gastrointestinal or bowel diseases in particular [14].

Our data suggested an indication of possible association between polymorphisms in the *hMSH6* gene and risk of CRC. The increased risk of CRC and, in particular, of rectal cancer, in association with the variant allele of the -556G > T polymorphism in the *hMSH6* promoter may be due to its modulating effect on transcription. An *in vitro* study on CHO cells suggested that the same sequence variation results in loss of Sp1 binding site involved in the gene transcription regulation. In addition, -556T allele in combination with two other polymorphisms in *hMSH6* (-448G > A and -159C > T), apparently affects gene expression by promoter methylation [8]. However, no association of the -159C > T polymorphism with CRC risk was reported in an independent study [15]. We also found that the carriers of the variant allele for the -145G > A polymorphism in the *hMSH6* exon 1 coding for Gly39Glu amino acid change exhibited a significantly lower risk of CRC than non-carriers. This tendency was observed in patients with both colon and rectal cancers. Previous studies did not find any association of the same polymorphism with CRC risk [16] and with risk of adenomatous polyposis in populations of Caucasian origin [17]. Glycine-to-glutamic acid substitution can determine the formation of sterically different helical structures, polypeptide folding, and intrinsic aggregation due to hydrophilic side chain of glutamic acid, thus affecting the protein function [18,19].

We found that the -93G > A SNP located in the core promoter region of the *hMLH1* gene had no effect on CRC susceptibility, even though it has been demonstrated earlier to modulate the risk of lung [20], breast [21], and endometrial cancers [22]. The -93G > A polymorphism is located in a CpG island and the A-allele is presumably involved in the epigenetic silencing of *hMLH1*, via promoter methylation [9]. Recently, the presence of variant allele was significantly associated with increased risk of MSI-unstable CRC [23]. One of the plausible reasons for not finding any association could be that patients with CRC were not classified according to MSI status in our study population. In general, MSI-unstable tumors account for 15–20% of sporadic CRC malignancies [24].

In an earlier study, the intronic substitution IVS12-6T > C in the *hMSH2* gene was associated with predisposition to sporadic CRC in a small European study [25]. However, our data did not show any association for this polymorphism with the risk of the disease, in concordance with studies carried out on a Korean population [26] and on a population with mixed ethnicity from Canada [23]. The coding polymorphism Ex6 + 23G > A in the *hMSH2* gene may have a putative influence on MMR function due to location in a region stabilizing this gene's interaction with hEXO1 protein [27] and was considered as affecting breast cancer susceptibility [13], with no significant association with CRC risk [28].

Few studies are available at present for other SNPs investigated in the remaining genes, *hMSH3* and *hEXO1*, for which we found no significant associations. An association with increased CRC risk has been shown for Ex23 + 3G > A polymorphism in *hMSH3* in patients of European descent [15,29] and Ex12 + 49C > T SNP in *hEXO1* in a Japanese population [30]. Both polymorphisms could have a putative functional effect since Ex23 + 3G > A cause Ala1045Thr amino acid change in ATP-binding domain of *hMSH3* and Ex12 + 49C > T coding Thr439Met substitution in *hEXO1* may influence interaction between hEXO1 and hMLH1 proteins [27]. Although other polymorphisms in other MMR genes, such as *hMLH3*, *hPMS1*, and *hPMS2* could be of relevance to the CRC risk [28,31–33], these were not addressed in the current study.

Haplotype analysis represents a powerful approach in searching for disease-causing alleles and in locating the underlying mutations. We found that the haplotype TAG, based on three SNPs of the *hMSH6* gene (-556G > T-Ex1 – 145G > A-IVS4-101G > C), was associated with a decreased overall risk of CRC. We also observed that the most frequent haplotype, GGG, was associated with an increased risk of rectal cancer.

No unambiguous agreement could be presently reached about the role of haplotypes in *hMLH1*, *hMSH2*, or *hMSH3* genes. However, for none of these genes we recorded a significant association with sporadic CRC risk. The increased CRC risk was proposed for a rare *hMLH1* haplotype [29] and for *hMSH3* haplotype constructed using two SNPs in the gene including Ex23 + 3G > A [16].

To conclude, in our analysis of modulating effects of individual MMR polymorphisms on sporadic CRC risk we found an indication of an association with polymorphisms and haplotypes in the *hMSH6* gene. Though, after correction for multiple hypotheses test-

ing our results cannot be considered as statistically significant and, thus, should not be overinterpreted. This is in line with the results of recent studies describing a limited role of the common variants in the MMR genes *hMLH1*, *hMSH2*, *hMSH3* and *hMSH6* for CRC risk [28,29]. Nevertheless, considering the importance of the MMR genes in the etiology of CRC further studies with pooled data would be imperative to determine if the common variants in the genes *per se* or in combination with other variants play any role in the disease pre-disposition.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was supported by the grants: GACR 310/07/1430, AVOZ 50390703 and 50390512 of the Czech Republic and by BMBF for NGFN+.

References

[1] R.R. Iyer, A. Pluciennik, V. Burdett, P.L. Modrich, DNA mismatch repair: functions and mechanisms, Chem. Rev. 106 (2006) 302–323.
[2] G.M. Li, Mechanisms and functions of DNA mismatch repair, Cell Res. 18 (2008) 85–98.
[3] D.L. Worthley, V.L. Whitehall, K.J. Spring, B.A. Leggett, Colorectal carcinogenesis: road maps to cancer, World J. Gastroenterol. 13 (2007) 3784–3791.
[4] R.A. Barnetson, A. Tenesa, S.M. Farrington, I.D. Nicholl, R. Cetnarskyj, M.E. Porteous, H. Campbell, M.G. Dunlop, Identification and survival of carriers of mutations in DNA mismatch-repair genes in colon cancer, N. Engl. J. Med. 354 (2006) 2751–2763.
[5] I. Tomlinson, E. Webb, L. Carvajal-Carmona, P. Broderick, Z. Kemp, S. Spain, S. Penegar, I. Chandler, M. Gorman, W. Wood, E. Barclay, S. Lubbe, L. Martin, G. Sellick, E. Jaeger, R. Hubner, R. Wild, A. Rowan, S. Fielding, K. Howarth, A. Silver, W. Atkin, K. Muir, R. Logan, D. Kerr, E. Johnstone, O. Sieber, R. Gray, H. Thomas, J. Peto, J.B. Cazier, R.A. Houlston, A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21, Nat. Genet. 39 (2007) 984–988 (CORGI Consortium).
[6] I. Tomlinson, E. Webb, L. Carvajal-Carmona, P. Broderick, K. Howarth, A.M. Pittman, S. Spain, S. Lubbe, A. Walther, K. Sullivan, E. Jaeger, S. Fielding, A. Rowan, J. Vijayakrishnan, E. Domingo, I. Chandler, Z. Kemp, M. Qureshi, S.M. Farrington, A. Tenesa, J.G.D. Prendergast, R.A. Barnetson, S. Penegar, E. Barclay, W. Wood, L. Martin, M. Gorman, H. Thomas, J. Peto, D.T. Bishop, R. Gray, E.R. Maher, A. Lucassen, D. Kerr, D.G.R. Evans, C. Schafmayer, S. Buch, H. Volzke, J. Hampe, S. Schreiber, U. John, T. Koessler, P. Pharoah, T. van Wezel, H. Morreau, J.T. Wijnen, J.L. Hopper, M.C. Southey, G.G. Giles, G. Severi, S. Castellvı-Bel, C. Ruiz-Ponte, A. Carracedo, A. Castells, A. Forsti, K. Hemminki, P. Vodicka, A. Naccarati, L. Lipton, J.W. Ho, K.K. Cheng, P.C. Sham, J. Luk, J.A. Agúndez, J.M. Ladero, M. de la Hoya, T. Caldés, I. Niittymäki, S. Tuupanen, A. Karhu, L. Aaltonen, J.B. Cazier, H. Campbell, M.G. Dunlop, R.S. Houlston, A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14, 8q23.3, Nat. Genet. 40 (2008) 623–630 (The CORGI Consortium).
[7] J. Jiricny, M. Nystrom-Lahti, Mismatch repair in cancer, Curr. Opin. Genet. Dev. 10 (2000) 157–161.
[8] I. Gazzoli, R.D. Kolodner, Regulation of the human *MSH6* gene by the Sp1 transcription factor and alteration of promoter activity and expression by polymorphisms, Mol. Cell Biol. 23 (2003) 7992–8007.
[9] H. Chen, N.P. Taylor, K.M. Sotamaa, D.G. Mutch, M.A. Powell, A.P. Schmidt, S. Feng, H.L. Hampel, A. de la Chapelle, P.J. Goodfellow, Evidence for heritable predisposition to epigenetic silencing of MLH1, Int. J. Cancer 120 (2007) 1684–1688.
[10] B.W. Stewart, P. Kleihues (Eds.), World Cancer Report, IARC Press, Lyon, 2003, pp. 198–202.
[11] D.M. Parkin, F. Bray, J. Ferlay, P. Pisani, Global cancer statistics, 2002, CA Cancer J. Clin. 55 (2005) 74–108.
[12] D. Landi, F. Gemignani, A. Naccarati, B. Pardini, P. Vodicka, L. Vodickova, J. Novotny, A. Försti, K. Hemminki, F. Canzian, S. Landi, Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer, Carcinogenesis 29 (2008) 579–584.
[13] T. Poplawski, M. Zadrozny, A. Kolacinska, J. Rykala, Z. Morawiec, J. Blasiak, Polymorphisms of the DNA mismatch repair gene *HMSH2* in breast cancer occurence and progression, Breast Cancer Res. Treat. 94 (2005) 199–204.
[14] M.E. Martinez, J.R. Marshall, E. Giovannucci. Diet and cancer prevention: the roles of observation and experimentation, Nature Reviews Cancer (2008), advance online publication August 7.
[15] M. Mrkonjic, S. Raptis, R.C. Green, N. Monga, D. Daftary, E. Dicks, H.B. Younghusband, P.S. Parfrey, S.S. Gallinger, J.R. McLaughlin, J.A. Knight, B. Bapat, MSH2 118T > C and MSH6 159C > T promoter polymorphisms and the risk of colorectal cancer, Carcinogenesis 28 (2007) 2575–2580.
[16] S.I. Berndt, E.A. Platz, M.D. Fallin, L.W. Thuita, S.C. Hoffman, K.J. Helzlsouer, Mismatch repair polymorphisms and the risk of colorectal cancer, Int. J. Cancer 120 (2007) 1548–1554.
[17] J.H. Yu, J. Bigler, J. Whitton, J.D. Potter, C.M. Ulrich, Mismatch repair polymorphisms and colorectal polyps: hMLH1-93G > A variant modifies risk associated with smoking, Am. J. Gastroenterol. 101 (2006) 1313–1319.
[18] C. Branden, J. Tooze, Introduction to Protein Structure. Garland Publishing, 2nd edition, Taylor and Francis Group, New York, 1999.
[19] F. Chiti, M. Stefani, N. Taddei, G. Ramponi, C.M. Dobson, Rationalization of the effects of mutations on peptide and protein aggregation rates, Nature 424 (2003) 805–808.
[20] S.H. Park, G.Y. Lee, H.S. Jeon, S.J. Lee, K.M. Kim, S.S. Jang, C.H. Kim, W.K. Lee, S. Kam, R.W. Park, I.S. Kim, T.H. Jung, J.Y. Park, -93G-A polymorphism of hMLH1 and risk of primary lung cancer, Int. J. Cancer 112 (2004) 678–682.
[21] K.M. Lee, J.Y. Choi, C. Kang, C.P. Kang, S.K. Park, H. Cho, D.Y. Cho, K.Y. Yoo, D.Y. Noh, S.H. Ahn, C.G. Park, Q. Wei, D. Kang, Genetic polymorphisms of selected DNA repair genes, estrogen and progesterone receptor status, and breast cancer risk, Clin. Cancer Res. 11 (2005) 4620–4626.
[22] M.E. Beiner, B. Rosen, A. Fyles, I. Harley, T. Pal, K. Siminovitch, S. Zhang, P. Sun, S.A. Narod, Endometrial cancer risk is associated with variants of the mismatch repair genes *MLH1* and *MSH2*, Cancer Epidemiol. Biomarkers Prev. 15 (2006) 1636–1640.
[23] S. Raptis, M. Mrkonjic, R.C. Green, V.V. Pethe, N. Monga, Y.M. Chan, D. Daftary, E. Dicks, B.H. Younghusband, P.S. Parfrey, S.S. Gallinger, J.R. McLaughlin, J.A. Knight, B. Bapat, MLH1-93G > A promoter polymorphism and the risk of microsatellite-unstable colorectal cancer, J. Natl. Cancer Inst. 99 (2007) 463–474.
[24] C.R. Boland, S.N. Thibodeau, S.R. Hamilton, D. Sidransky, J.R. Eshleman, R.W. Burt, S.J. Meltzer, M.A. Rodriguez-Bigas, R. Fodde, G.N. Ranzani, S. Srivastava, A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer, Cancer Res. 58 (1998) 5248–5257.
[25] C. Goessl, J. Plaschke, S. Pistorius, M. Hahn, S. Frank, M. Hampl, H. Görgens, R. Koch, H.-D. Saeger, H.K. Schackert, An intronic germline transition in the HNPCC gene *hMSH2* is associated with sporadic colorectal cancer, Eur. J. Cancer 33 (1997) 1869–1874.
[26] J.C. Kim, S.A. Roh, K.H. Koo, I.H. Ka, H.C. Kim, C.S. Yu, K.H. Lee, J.S. Kim, H.I. Lee, W.F. Bodmer, Genotyping possible polymorphic variants of human mismatch repair genes in healthy Korean individuals and sporadic colorectal cancer patients, Fam. Cancer 3 (2004) 129–137.
[27] C. Schmutte, M.M. Sadoff, K.-S. Shim, S. Acharya, R. Fishel, The interaction of DNA mismatch repair proteins with human exonuclease I, J. Biol. Chem. 276 (2001) 33011–33018.
[28] C. Schafmayer, S. Buch, J.H. Egberts, A. Franke, M. Brosch, A. El Sharawy, M. Conring, M. Koschnick, S. Schwiedernoch, A. Katalinic, B. Kremer, U.R. Fölsch, M. Krawczak, F. Fändrich, S. Schreiber, J. Tepel, J. Hampe, Genetic investigation of DNA-repair pathway genes *PMS2*, *MLH1*, *MSH2*, *MSH6*, *MUTYH*, *OGG1* and *MTH1* in sporadic colon cancer, Int. J. Cancer 121 (2007) 555–558.
[29] T. Koessler, M.Z. Oestergaard, H. Song, J. Tyrer, B. Perkins, A. Dunning, D. Easton, P.P. Pharoah, Common variants in mismatch repair genes and risk of colorectal cancer, Gut 57 (2008) 1097–1101.
[30] H. Yamamoto, H. Hanafusa, M. Ouchida, M. Yano, H. Suzuki, M. Murakami, M. Aoe, N. Shimizu, K. Nakachi, K. Shimizu, Single nucleotide polymorphisms in the *EXO1* gene and risk of colorectal cancer in a Japanese population, Carcinogenesis 26 (2005) 411–416.
[31] Z.Q. Yuan, B. Gottlieb, L.K. Beitel, N. Wong, P.H. Gordon, Q. Wang, A. Puisieux, W.D. Foulkes, M. Trifiro, Polymorphisms and HNPCC: PMS2–MLH1 protein interactions diminished by single nucleotide polymorphisms, Hum. Mutat. 19 (2002) 108–113.
[32] A. Loukola, S. Vilkki, J. Singh, V. Launonen, L.A. Aaltonen, Germline and somatic mutation analysis of *MLH3* in MSI-positive colorectal cancer, Am. J. Pathol. 157 (2000) 347–352.
[33] J.A. Heck, J.L. Argueso, Z. Gemici, R.G. Reeves, A. Bernard, C.F. Aquadro, E. Alani, Negative epistasis between natural variants of the *Saccharomyces cerevisiae* *MLH1* and *PMS1* genes results in a defect in mismatch repair, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 3256–3261.