Impact of Genetic Variation in MicroRNA-binding Site on Susceptibility to Colorectal Cancer

BYUNG WOOG KANG[1], HYO-SUNG JEON[2*], YEE SOO CHAE[1], SOO JUNG LEE[1], JUN SEOK PARK[3], GYU SEOG CHOI[3] and JONG GWANG KIM[1*]

[1]*Department of Oncology/Hematology, Kyungpook National University Medical Center, Kyungpook National University School of Medicine, Cancer Research Institute, Daegu, Republic of Korea;*
[2]*M monitor Inc., Daegu, Republic of Korea;*
[3]*Department of Surgery, Kyungpook National University Medical Center, Kyungpook National University School of Medicine, Daegu, Republic of Korea*

Abstract. *The present study analyzed single nucleotide polymorphisms (SNPs) located at putative microRNA(miRNA)-binding sites of the 3'-untranslated region (UTR) in different genes and investigated their impact on the susceptibility to colorectal cancer (CRC). Ninety-two SNPs were selected using an in silico analysis of 3'-UTR SNPs in an SNP database and their miRNA binding efficiency was calculated using several miRNA databases and the HapMap database. Two independent study sets were used: 380 healthy controls and 371 patients with colorectal adenocarcinoma for the discovery set, and 521 healthy controls and 524 patients with colorectal adenocarcinoma for the validation set. The SNP genotyping was performed using a Sequenom MassARRAY. In addition, a luciferase assay was used to investigate whether miR-370 modulated docking protein 3 (DOK3) gene expression when rs2279398G>A was included in the DOK3 3'-UTR region. For the discovery set, 16 out of 92 SNPs were significantly associated with the risk of CRC in at least one of the genetic models. The validation set showed that among these 16 SNPs, DOK3 rs2279398G>A was significantly associated with reduced risk of CRC in a recessive model [adjusted odds ratio (aOR)=0.65, 95% confidence interval (CI)=0.44-0.97, p=0.03]. In a combined analysis, DOK3 rs2279398G>A was associated with a significantly reduced risk of CRC in a co-dominant and recessive model (aOR=0.84, 95% CI=0.73-0.96, p=0.012; aOR=0.65, CI=0.49-0.88, p=0.004, respectively). Significantly lower Renilla activity was also observed with the rs2279398 AA construct when compared to the rs2279398 GG construct (p<0.001). DOK3 rs2279398G>A may affect the expression of DOK3 by altering the miRNA binding efficiency at the miRNA-binding sites of the 3'-UTR in DOK3, thereby impacting CRC tumorigenesis.*

Colorectal cancer (CRC) is the third most common form of cancer and the fourth most common cause of cancer deaths globally (1). Since relapse following cancer surgery and resistance to chemotherapy persist as major obstacles to effective treatment, prognosis for CRC remains poor. Therefore, the rising incidence of CRC and its poor prognosis are pushing systematic approaches towards a more precise identification of patients with a higher susceptibility of CRC (2). While genetic factors play an important role and provide valuable information on the etiology of both sporadic and familial CRC, fewer than 6% of CRC cases can be explained by rare, high-penetrance variants of the CRC susceptibility genes identified to date (3, 4).

MicroRNAs (miRNAs) are a class of short, single-strand noncoding RNAs that play key roles in the regulation of gene expression by translational repression or mRNA degradation of the target, affecting critical functions in various physiological processes (5, 6). Alterations of miRNAs are involved in a wide range of human diseases, and can also affect tumor development and progression by regulating the expression of proto-oncogenes and tumor-suppressor genes (7, 8). Recent reports have shown that the expression of miRNAs is associated with the development

*These Authors contributed equally to this work.

Correspondence to: Jong Gwang Kim, MD, Ph.D., Department of Oncology/Hematology, Kyungpook National University Medical Center, Kyungpook National University School of Medicine, Cancer Research Institute, 807 Hogukno, Buk-Gu, Daegu, 702-210, Korea. Tel: +82 532003521, Fax: +82 532002029, e-mail: jkk21c@knu.ac.kr and Hyo-Sung Jeon, Ph.D., Department of Biochemistry and Cell Biology, Kyungpook National University School of Medicine, 807 Hogukno, Buk-gu, Daegu, 702-210, Korea. Tel: +82 1083318014, Fax: +82 532002029, e-mail: jeonh@knu.ac.kr

Key Words: Colorectal cancer, microRNA target gene, polymorphism, susceptibility, DOK3.

or prognosis of CRC (9-11). Therefore, miRNAs are likely to be highly useful as biomarkers in CRC. Single nucleotide polymorphisms (SNPs) can occur in genes for the miRNA biogenesis pathway, primary miRNA, pre-miRNA, or mature miRNA sequences (12). Although the exact role of SNPs detected in miRNA has yet to be clarified, our previous studies demonstrated that specific SNPs are associated with susceptibility to CRC and its prognosis (13-15). Importantly, the predominant function of miRNAs is to regulate protein translation by binding to complementary sequences in the 3'-untranslated region (UTR) of target mRNAs, and thereby negatively regulate mRNA translation (16). Thus, polymorphisms of miRNA-binding sites in the 3'-UTRs of target genes could potentially affect the efficiency of miRNA binding to the target sites by altering the miRNA–mRNA interaction and changing the expression of the miRNA targets (12, 17, 18). However, these SNPs and their relationship to the risk of CRC have not yet been fully characterized as novel biomarkers for categorizing patients with different genetic risks (19). Accordingly, the present study analyzed 92 SNPs located at putative miRNA-binding sites in the 3'-UTR of different genes and investigated their impact on susceptibility to CRC in terms of altering the miRNA-binding efficiency.

Patients and Methods

Study populations. A total of 371 patients who were histologically confirmed to have sporadic CRC at the Kyungpook National University Hospital (KNUH) in Daegu, Korea, were included in the discovery set. The controls (n=380) were randomly selected from a pool of healthy volunteers who visited the general health check-up center at the hospital and were matched with the patients as regards gender. For an independent validation set, a total of 524 patients and 521 controls were collected. This study was approved by the Institutional Review Board of KNUH (KNUMCBIO_11-1002) and written informed consent was obtained from all participants. The genomic DNA samples from the patients with CRC and healthy controls were provided by the National Biobank of Korea-KNUH, which is supported by the Korean Ministry of Health, Welfare and Family Affairs. All patients and controls were ethnic Koreans residing in Daegu or the surrounding regions. The CRC diagnosis and staging were determined according to the WHO classifications and TMN classifications included in the 6th edition of the American Joint Committee on Cancer (20).

SNP selection and genotyping. A total of 12,877 polymorphisms in the 3'-UTR of genes that may be affected by miRNAs were selected from a public SNP database (http://www.ncbi.nlm.nih.gov/SNP) and Ensemble data. To select all the potentially functional miRNA-binding polymorphisms, this study used the miRNA target prediction program miRanda algorithm version 3.1 (http://www.microrna.org/microrna), where SNPs with minimum free energy (MFE) difference <6 between two alleles were excluded when calculating the MFE for all the possible miRNA and SNP complex sequences. As a result, a total of 391 SNPs were collected with a minor allele frequency (MAF) ≥0.1 from the HapMap JPT data in the public SNP database and a Gibbs binding free energy (ΔG, kJ/mol) ≥5.7 kJ/mol in a miRNA and RNA hybrid model. The selected 391 SNPs were analyzed with a healthy control population using SEQUENOM's MassARRAY® iPLEX assay (SEQUENOM Inc., San Diego, CA, USA), which eliminated 54 SNPs with a MAF <0.1. Among the remaining 337 SNPs, 92 SNPs applicable to the SEQUENOM's MassARRAY were genotyped.

The genomic DNA was extracted from normal tissues using as QIAamp® genomic DNA kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The genotype analysis was performed using SEQUENOM's MassARRAY® iPLEX assay according to the instructions of the manufacturer. To validate the genotyping, approximately 5% of the cohort samples were randomly selected for re-genotyping by another researcher using a sequencing method or a restriction fragment length polymorphism assay, and the results were 100% concordant.

Cloning of luciferase reporter gene and luciferase assay. In the in silico analysis, since docking protein 3 (DOK3) gene expression was affected by miR-370-binding efficiency, miR-370 was selected using the MicroSNIPer program (http://epicenter.ie-freiburg.mpg.de/services/microsniper/index.php). A luciferase assay was used to investigate whether *miR-370* modulated *DOK3* gene expression when rs2279398 was included in the *DOK3* 3'-UTR region. A 142-bp fragment of the *DOK3* 3'-UTR region was synthesized using polymerase chain reaction and cloned into the dual luciferase vector psiCHECK2 (Promega, Madison, WI, USA). A forward primer with an *Xho*I restriction site (5'- CCC CTC GAG ACA CAC GCC TGT GTC CAC C -3') and reverse primer with a *Not*I restriction site (5'-CCG CGG CCG CTG CAG TGG GTT TGA GC -3') were then used to amplify the *DOK3* 3'-UTR region. DNA sequencing was also used to verify that the clone sequences were all correct. Using a 12-well plate, 293T cells (ATCC, Manassas, VA, USA) were seeded in a Dulbeco's modified Eagle's medium supplemented with 10% heat-inactivated fetal bovine serum. The cells were then transfected with psiCHECK2-*DOK3* constructs containing the 3'-UTR of *DOK3* that included rs2279398G>A, in the presence of a *miR-370* mimic (Qiagen) using Effectene (Qiagen) transfection reagent. After 48 hours, the cells were collected and cell lysates prepared according to Promega's instruction manual. The Renilla luciferase activity was measured using a Lumat LB953 luminometer (EG & G Berthhold, Bad Wildbad, Germany), and the results normalized using the luciferase activity. All the experiments were performed in triplicate.

Statistical analysis. The results for the patients and controls were compared using Student's *t*-test and the chi-square test for the continuous variables and categorical variables, respectively. The *p*-value for the luciferase assay was determined using Student's *t*-test. The Hardy–Weinberg equilibrium was tested by comparing the observed and expected genotype frequencies using a χ^2 test with 1 degree of freedom. An unconditional logistic regression analysis was used to calculate the odds ratios (ORs) and 95% confidence intervals (CIs), with adjustment for age and gender. All the analyses were performed using Statistical Analysis Software for Windows, version 9.2 (SAS Institute, Cary, NC, USA). The statistical significance of difference between the groups was determined using Student's t-test with the treatment and experiment as factors. Differences were considered statistically significant at $p<0.05$.

Table I. *Characteristics of the study population.*

		Discovery set			Validation set		
		Cases n (%)	Controls n (%)	*p*-Value*	Cases n (%)	Controls n (%)	*p*-Value*
Total		371	380		524	521	
Gender	Male	236 (63.6)	252 (66.3)	0.437	279 (53.2)	280 (53.7)	0.872
	Female	135 (36.4)	128 (33.7)		245 (46.8)	241 (46.3)	
Age, years	Mean (range)	60.9 (30-87)	59.7 (26-84)	0.006	62.4 (24-86)	62.0 (28-86)	0.742
	<61	201 (54.2)	168 (44.2)		309 (59.0)	302 (58.0)	
	≥61	170 (45.8)	212 (55.8)		215 (41.0)	219 (42.0)	
Tumor site	Colon	198(53.4)			307 (58.6)		
	Rectum	170 (45.8)			215 (41.0)		
	Other	3 (0.8)			2 (0.4)		
Stage	I	66 (17.8)			82 (15.6)		
	II	113 (30.5)			214 (40.8)		
	III	176 (47.4)			153 (29.2)		
	IV	16 (4.3)			75 (14.3)		
CEA, unit (ng/mL)	Mean (range)	7.5 (0.1-193.0)			14.4 (0.1-1042.0)		
	<7	75 (20.2)			102 (19.5)		
	≥7	296 (79.8)			422 (80.5)		

CEA: Carcinoembryonic antigen. *Log-rank test.

Results

Characteristics of study populations. The descriptive characteristics of the study participants are provided in Table I. No significant differences were found between the patients and the controls as regards the mean age and sex in the discovery set and validation set, yet the mean age of the patients in the discovery set was significantly higher than that of the controls (p=0.006). The median age of patients with CRC was 65 years, and 505 (56.4%) had colonic cancer and 385 (43.0%) had rectal cancer. The pathological stages after surgical resection were as follows: stage I: n=148, 16.5%; stage II: n=327, 36.5%; stage III: n=329, 36.8%; and stage IV: n=91, 10.2%. Twenty percent of the patients had elevated carcinoembryonic antigen levels.

Genotypic frequencies and their association with CRC risk. For the discovery set, 16 of the 92 SNP genotypes were significantly associated with the risk of CRC in at least one of the genetic models (Table II). Meanwhile, for the validation set, among these 16 SNPs, *DOK3* rs2279398G>A was significantly associated with reduced CRC risk in a recessive model [adjusted odds ratio (aOR)=0.65, 95% CI=0.44-0.97, p=0.033]. In the combined analysis, the A allele of *DOK3* rs2279398G>A was associated with a significantly reduced risk of CRC co-dominant and recessive models (aOR=0.84, 95% CI=0.73-0.96, p=0.012, aOR=0.65, CI=0.49-0.88, p=0.004, respectively) (Table III). Finally, the subgroup analysis according to the primary tumor site also revealed a strong association of the A allele of *DOK3* rs2279398G>A with reduced risk of colonic cancer (Table IV).

Effect of DOK3 rs2279398G>A polymorphism on miR-370 binding. To verify whether the rs2279398G>A polymorphism directly affected the miR-370-binding efficiency in *DOK3* 3'-UTR, a Renilla luciferase reporter plasmid cloned downstream to a segment of the *DOK3* 3'-UTR containing rs2279398G>A was generated. Each construct, psiCHECK2: rs2279398GG and psiCHECK2: rs2279398AA, was then co-transfected into 293T cells with *miR-370*, and the Renilla luciferase activity measured after 48 hours. As shown in Figure 1, significantly lower Renilla activity was observed using the rs2279398AA construct when compared to the rs2279398GG construct (p<0.001), suggesting that the rs2279398G>A polymorphism may affect *DOK3* expression by the direct binding of miR-370 to its 3'-UTR.

Discussion

A relatively large population-based study was used to investigate the influence of 92 SNPs located at putative miRNA-binding sites of the 3'-UTR of different genes on the risk of CRC. As a result, the *DOK3* rs2279398G>A polymorphism was found to be associated with a significantly lower risk of CRC under a co-dominant and recessive model for the minor A allele. Therefore, these findings suggests that the polymorphism at the miRNA-binding site of *DOK3* (rs2279398) may be involved in CRC carcinogenesis and

Table II. *Genotypes in the discovery and validation studies, and their association with the risk of colorectal cancer.*

	Discovery		Validation		Combined	
	aOR (95% CI)	*p*-Value	aOR (95% CI)	*p*-Value	aOR (95% CI)	*p*-Value
PCSK4 rs1047854 (C>T)						
CC	1.00		1.00		1.00	
CT	0.85 (0.61-1.19)	0.342	1.30 (0.99-1.70)	0.061	1.09 (0.89-1.35)	0.405
TT	0.63 (0.41-0.96)	0.033	1.38 (0.95-1.99)	0.091	0.98 (0.74-1.29)	0.887
Dominant	0.78 (0.57-1.07)	0.121	1.32 (1.02-1.71)	0.037	1.06 (0.87-1.29)	0.550
Recessive	0.69 (0.47-1.01)	0.056	1.18 (0.85-1.65)	0.327	0.93 (0.73-1.19)	0.571
Co-dominant	0.80 (0.65-0.99)	0.037	1.20 (1.00-1.43)	0.049	1.01 (0.88-1.15)	0.919
FAM151A rs11206394 (G>C)						
GG	1.00		1.00		1.00	
GC	0.77 (0.57-1.05)	0.100	1.04 (0.81-1.35)	0.750	0.92 (0.75-1.12)	0.378
CC	0.30 (0.15-0.59)	0.001	0.94 (0.57-1.54)	0.800	0.60 (0.41-0.89)	0.011
Dominant	0.68 (0.51-0.92)	0.011	1.03 (0.80-1.31)	0.838	0.86 (0.71-1.04)	0.117
Recessive	0.33 (0.17-0.65)	0.001	0.92 (0.57-1.508)	0.743	0.62 (0.42-0.91)	0.016
Co-dominant	0.66 (0.51-0.84)	0.001	1.00 (0.82-1.22)	0.975	0.84 (0.72-0.98)	0.025
RAB24 rs1128287 (C>A)						
CC	1.00		1.00		1.00	
CA	0.65 (0.46-0.92)	0.015	0.93 (0.70-1.23)	0.599	0.80 (0.65-1.00)	0.051
AA	0.73 (0.48-1.09)	0.125	0.99 (0.70-1.41)	0.967	0.87 (0.67-1.14)	0.312
Dominant	0.68 (0.49-0.93)	0.017	0.95 (0.72-1.24)	0.683	0.82 (0.67-1.01)	0.067
Recessive	0.95 (0.67-1.34)	0.772	1.04 (0.77-1.41)	0.784	1.00 (0.80-1.26)	0.989
Co-dominant	0.84 (0.69-1.03)	0.092	0.99 (0.83-1.18)	0.917	0.92 (0.81-1.06)	0.243
IL16 rs1131445 (T>C)						
TT	1.00		1.00		1.00	
TC	1.22 (0.90-1.66)	0.200	1.40 (1.08-1.81)	0.010	1.33 (1.09-1.62)	0.004
CC	0.60 (0.35-1.04)	0.069	0.97 (0.60-1.57)	0.893	0.79 (0.55-1.13)	0.201
Dominant	1.09 (0.81-1.45)	0.584	1.33 (1.04-1.70)	0.025	1.23 (1.02-1.48)	0.034
Recessive	0.55 (0.32-0.93)	0.025	0.83 (0.52-1.32)	0.428	0.69 (0.49-0.98)	0.038
Co-dominant	0.94 (0.75-1.17)	0.567	1.16 (0.95-1.41)	0.146	1.06 (0.92-1.23)	0.436
TPRX1 rs12462695 (C>A)						
CC	1.00		1.00		1.00	
CA	0.93 (0.68-1.27)	0.638	1.01 (0.78-1.31)	0.934	0.98 (0.80-1.19)	0.833
AA	0.58 (0.35-0.96)	0.032	1.31 (0.86-1.97)	0.206	0.94 (0.68-1.29)	0.694
Dominant	0.85 (0.63-1.14)	0.269	1.06 (0.83-1.36)	0.619	0.97 (0.80-1.17)	0.757
Recessive	0.60 (0.37-0.97)	0.036	1.30 (0.88-1.92)	0.193	0.95 (0.70-1.28)	0.730
Co-dominant	0.82 (0.65-1.02)	0.070	1.10 (0.91-1.32)	0.327	0.97 (0.84-1.12)	0.694
RGS5 rs15049 (C>A)						
CC	1.00		1.00		1.00	
CA	0.72 (0.53-0.98)	0.037	0.94 (0.73-1.21)	0.625	0.85 (0.70-1.03)	0.099
AA	0.47 (0.28-0.80)	0.005	0.85 (0.54-1.36)	0.508	0.66 (0.47-0.94)	0.021
Dominant	0.67 (0.50-0.90)	0.007	0.92 (0.72-1.18)	0.524	0.81 (0.67-0.98)	0.030
Recessive	0.55 (0.33-0.91)	0.021	0.88 (0.56-1.38)	0.574	0.72 (0.51-1.00)	0.051
Co-dominant	0.70 (0.56-0.88)	0.002	0.93 (0.77-1.13)	0.460	0.83 (0.72-0.96)	0.011
C1orf177 rs17111100 (C>A)						
CC	1.00		1.00		1.00	
CA	1.37 (0.99-1.90)	0.062	1.13 (0.85-1.50)	0.394	1.23 (1.00-1.53)	0.056
AA	1.37 (0.89-2.12)	0.155	1.28 (0.90-1.82)	0.165	1.33 (1.02-1.75)	0.039
Dominant	1.37 (1.00-1.87)	0.049	1.17 (0.90-1.53)	0.242	1.26 (1.03-1.54)	0.026
Recessive	1.13 (0.77-1.66)	0.534	1.19 (0.88-1.61)	0.266	1.17 (0.92-1.48)	0.195
Co-dominant	1.20 (0.97-1.48)	0.093	1.13 (0.95-1.35)	0.162	1.16 (1.02-1.33)	0.027
CCDC17 rs2275085 (T>C)						
TT	1.00		1.00		1.00	
TC	1.47 (1.08-1.99)	0.015	1.07 (0.83-1.39)	0.599	1.23 (1.01-1.50)	0.041
CC	0.18 (0.06-0.53)	0.002	1.16 (0.75-1.81)	0.501	0.81 (0.55-1.19)	0.279
Dominant	1.26 (0.94-1.70)	0.129	1.09 (0.85-1.39)	0.498	1.16 (0.96-1.40)	0.132
Recessive	0.15 (0.05-0.45)	0.001	1.13 (0.74-1.73)	0.578	0.74 (0.51-1.07)	0.111
Co-dominant	0.99 (0.77-1.29)	0.960	1.08 (0.89-1.30)	0.444	1.04 (0.90-1.21)	0.583

Table II. *Continued*

Table II. *Continued*

	Discovery		Validation		Combined	
	aOR (95% CI)	*p*-Value	aOR (95% CI)	*p*-Value	aOR (95% CI)	*p*-Value
DOK3 rs2279398 (G>A)						
GG	1.00		1.00		1.00	
GA	0.85 (0.62-1.17)	0.322	1.02 (0.78-1.32)	0.908	0.94 (0.77-1.15)	0.534
AA	0.60 (0.38-0.97)	0.036	0.66 (0.44-0.99)	0.044	0.63 (0.46-0.86)	0.004
Dominant	0.79 (0.59-1.06)	0.111	0.92 (0.72-1.18)	0.517	0.86 (0.71-1.04)	0.112
Recessive	0.65 (0.42-1.02)	0.060	0.65 (0.44-0.97)	0.033	0.65 (0.49-0.88)	0.004
Co-dominant	0.80 (0.65-0.99)	0.039	0.87 (0.73-1.05)	0.140	0.84 (0.73-0.96)	0.012
CYP4F8 rs2283606 (T>C)						
TT	1.00		1.00		1.00	
TC	0.58 (0.41-0.83)	0.002	0.85 (0.63-1.13)	0.262	0.73 (0.58-0.91)	0.005
CC	0.57 (0.21-1.61)	0.290	0.67 (0.32-1.43)	0.299	0.64 (0.35-1.17)	0.142
Dominant	0.58	(0.41-0.82)	0.002		0.83 (0.63-1.09)	0.179
0.72 (0.58-0.89)	0.002					
Recessive	0.65 (0.24-1.82)	0.417	0.70 (0.33-1.48)	0.347	0.69 (0.37-1.26)	0.221
Co-dominant	0.63 (0.46-0.85)	0.003	0.84 (0.66-1.06)	0.145	0.75 (0.62-0.90)	0.002
TSGA2 rs2839531 (T>C)						
TT	1.00		1.00		1.00	
TC	1.62 (1.17-2.25)	0.004	0.88 (0.68-1.15)	0.349	1.11 (0.91-1.36)	0.308
CC	1.84 (0.90-3.76)	0.096	1.18 (0.64-2.19)	0.591	1.41 (0.89-2.24)	0.147
Dominant	1.65 (1.21-2.26)	0.002	0.91 (0.71-1.17)	0.474	1.14 (0.94-1.39)	0.178
Recessive	1.55 (0.77-3.14)	0.224	1.24 (0.67-2.27)	0.492	1.36 (0.86-2.15)	0.191
Co-dominant	1.50 (1.15-1.95)	0.003	0.96 (0.78-1.19)	0.717	1.14 (0.97-1.35)	0.110
NOD2 rs3135499 (C>A)						
CC	1.00		1.00		1.00	
CA	1.45 (1.07-1.99)	0.018	0.97 (0.75-1.25)	0.800	1.14 (0.93-1.39)	0.199
AA	2.22 (1.33-3.70)	0.002	1.37 (0.82-2.31)	0.234	1.70 (1.183-2.447)	0.004
Dominant	1.59 (1.18-2.12)	0.002	1.01 (0.79-1.30)	0.912	1.216 (1.01-1.47)	0.042
Recessive	1.88 (1.15-3.08)	0.012	1.39 (0.84-2.32)	0.203	1.61 (1.13-2.29)	0.008
Co-dominant	1.48 (1.18-1.84)	0.001	1.06 (0.87-1.30)	0.551	1.23 (1.06-1.42)	0.007
TPST1 rs3757417 (G>T)						
GG	1.00		1.00		1.00	
GT	0.60 (0.43-0.83)	0.002	0.87 (0.66-1.15)	0.325	0.74 (0.60-0.92)	0.006
TT	0.83 (0.35-1.95)	0.663	0.82 (0.39-1.75)	0.613	0.83 (0.47-1.47)	0.524
Dominant	0.62 (0.45-0.85)	0.003	0.86 (0.66-1.13)	0.290	0.75 (0.61-0.92)	0.006
Recessive	0.97 (0.41-2.27)	0.938	0.85 (0.40-1.81)	0.681	0.91 (0.52-1.59)	0.727
Co-dominant	0.69 (0.52-0.9)	0.009	0.88 (0.70-1.12)	0.294	0.80 (0.67-0.95)	0.012
LOH12CR1 rs3803098 (A>G)						
AA	1.00		1.00		1.00	
AG	0.65 (0.46-0.90)	0.010	1.05 (0.81-1.37)	0.718	0.86 (0.70-1.06)	0.158
GG	0.84 (0.42-1.65)	0.605	1.81 (0.85-3.86)	0.125	1.19 (0.72-1.96)	0.489
Dominant	0.67 (0.49-0.92)	0.012	1.10 (0.85-1.42)	0.476	0.89 (0.73-1.09)	0.265
Recessive	0.95 (0.49-1.87)	0.888	1.78 (0.84-3.78)	0.133	1.250 (0.76-2.05)	0.378
Co-dominant	0.76 (0.59-0.98)	0.037	1.14 (0.91-1.42)	0.272	0.95 (0.80-1.12)	0.517
WDR51A rs747343 (C>T)						
CC	1.00		1.00		1.00	
CT	1.24 (0.87-1.76)	0.238	1.07 (0.80-1.42)	0.662	1.13 (0.91-1.41)	0.273
TT	0.78 (0.51-1.19)	0.244	1.23 (0.87-1.74)	0.240	1.02 (0.78-1.33)	0.909
Dominant	1.08 (0.77-1.51)	0.657	1.12 (0.85-1.46)	0.432	1.10 (0.89-1.35)	0.394
Recessive	0.67 (0.47-0.96)	0.028	1.18 (0.88 1.59)	0.266	0.94 (0.75-1.18)	0.573
Co-dominant	0.89 (0.72-1.10)	0.297	1.11 (0.93-1.32)	0.250	1.01 (0.89-1.16)	0.836
GJB3 rs9118 (C>G)						
CC	1.00		1.00		1.00	
CG	0.79 (0.58-1.07)	0.126	1.10 (0.86-1.43)	0.447	0.96 (0.79-1.16)	0.657
GG	0.61 (0.37-1.00)	0.048	1.00 (0.65-1.54)	0.994	0.80 (0.58-1.11)	0.181
Dominant	0.74 (0.56-0.99)	0.045	1.09 (0.85-1.39)	0.511	0.93 (0.77-1.12)	0.412
Recessive	0.67 (0.42-1.08)	0.103	0.95 (0.63-1.44)	0.809	0.82 (0.60-1.12)	0.206
Co-dominant	0.78 (0.63-0.97)	0.025	1.04 (0.86-1.25)	0.692	0.92 (0.80-1.06)	0.231

aOR: Odds ratio adjusted for age and gender; CI: confidence interval.

could be useful as a marker for genetic susceptibility to CRC. This result is worthy of note, since as far as we are aware, there are no previous data on the association of a *DOK* family polymorphism with CRC risk.

The 3'-UTR contains miRNA response elements and plays a crucial role in gene expression by influencing the localization, stability, export, and translation efficiency of mRNA, contributing to gene regulation (5, 6). Therefore, a polymorphism in the 3'-UTR of the gene can create as well as destroy the miRNA-binding site and have a similar influence on an SNP located within the miRNA seed region (21). In particular, when considering that miRNAs play a pivotal role in CRC initiation and development, it is important to evaluate the CRC risk associated with such SNPs. Several studies have already investigated the functional effect of various polymorphisms on the miRNA recognition sites on target genes and their association with CRC risk (22). Landi *et al*. studied selected SNPs within putative miRNA sites of genes known to be involved in CRC, and found that two polymorphisms, namely rs17281995 in [cluster of differentiation 86 (*CD86*)] and rs1051690 in [insulin receptor (*INSR*)], were associated with an increased risk of CRC (23, 24). Moreover, another recent study found an association between the SNPs in [replication protein A2 (*RPA2*)] and [general transcription factor IIH subunit (*GTF2H1*)] of a 3'-UTR involved in a DNA-repair pathway and the risk of rectal cancer (25). However, such biological functions still need to be experimentally verified through a functional study.

Interestingly, there are several important differences between the data in the present study and that in previous CRC association studies. Firstly, given the homogeneous ethnic background of Korean patients, any potential confounding effect due to ethnicity is likely to be small in the current study. Moreover, the luciferase reporter assay revealed that the *DOK3* (rs2279398) polymorphism directly affected gene expression and the risk of cancer. In addition, this study included a relatively large number of patients.

The significant finding from the present study is an association between a *DOK3* (rs2279398) polymorphism and the risk of CRC. *DOK3* is an adaptor protein that functions in feedback loops to modulate tyrosine kinase signaling. The activation of protein-tyrosine kinases induces tyrosine phosphorylation of the target proteins, triggering molecular interactions of the proteins. Therefore, the alteration of *DOK3* can negatively regulate the rat sarcoma virus-extracellular signal-regulated kinase (RAS-ERK) pathway, which is a positive signaling cascade critical for cell activation downstream of protein-tyrosine kinases (26). In a recent *in vitro* study of the response of CRC cell lines to a combination of 5-fluorouracil and radiotherapy, the authors reported that *DOK3* blocked the complex formation of downstream molecules, thereby behaving as an inhibitor of the RAS signaling pathway (27). Therefore, given these results, the rs2279398 polymorphism located in the 3'-UTR of the *DOK3* gene may influence miRNAbinding efficiency and thus potentially modulate tumor suppression. In fact, a recent phase III trial demonstrated that the addition of cetuximab, a monoclonal antibody that targets the epidermal growth factor receptor (EGFR), to chemotherapy significantly increased survival in patients with metastatic CRC (28). However, an anti-EGFR treatment may not be optimal for CRC, and interpreting the exact therapeutic effect of such a treatment is apparently difficult in patients with a V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*) mutation. Therefore, the current results may also provide essential SNP data to explain the precise biological significance of cancer development, as well as treatment resistance in patients with CRC. Similarly, Zhang *et al*. have reported that the *KRAS* 3'-UTR polymorphism can predict the cetuximab responsiveness in patients with wild-type *KRAS* with CRC treated with cetuximab monotherapy (29). Yet despite increasing evidence that *DOK3* plays important roles in several aspects of tumorigenesis, it is still unclear whether the *DOK3* (rs2279398) polymorphism itself alters the protein expression. Notwithstanding, the current study provides explicit evidence that the rs2279398 polymorphism is associated with a significantly lower risk of CRC under a co-dominant and recessive model for the minor A allele.

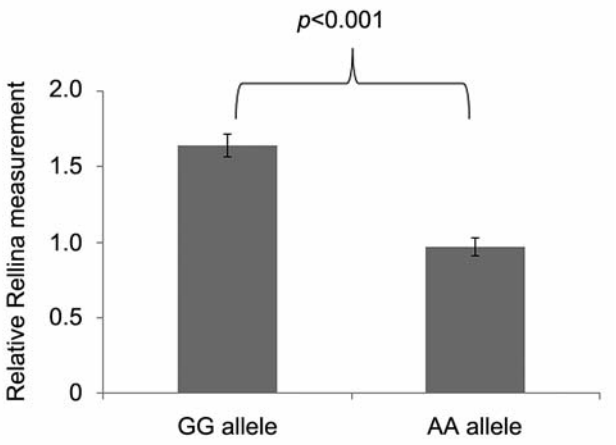

Figure 1. *Renilla/luciferase assay of miRNA suppression with docking protein 3 (DOK3) rs2279398G>A polymorphism. 293T cells were co-transfected with miR-370a and a psiCheck2 plasmid containing DOK3 3'-UTR with the rs279398GG or rs2279398AA polymorphism. Each bar represents the mean±s.e.m. renilla luciferase activity normalized to firefly luciferase activity. p-Value was determined using Student's t-test.*

To date, several risk factors have been associated with a higher incidence of CRC (30). The modifiable risk factors of CRC include smoking, physical inactivity, being overweight/obese, eating processed meat, and drinking alcohol excessively. The presence of polyps and hereditary

Table III. *Docking protein 3 (DOK3) genotypes of cases and controls, and their associations with the risk of colorectal cancer.*

	Discovery				Validation				Combined			
DOK3 rs2279398 genotype	Cases, n (%)	Controls, n (%)	aOR (95% CI)	*p*-Value	Cases, n (%)	Controls, n (%)	aOR (95% CI)	*p*-Value	Cases, n (%)	Controls, n (%)	aOR (95% CI)	*p*-Value
GG	184 (50.0)	158 (43.3)	1.00		255 (49.4)	237 (47.4)	1.00		439 (49.7)	395 (45.7)	1.00	
GA	147 (39.9)	153 (41.9)	0.85 (0.62-1.17)	0.322	213 (41.3)	195 (39.0)	1.02 (0.78-1.32)	0.908	360 (40.7)	348 (40.2)	0.94 (0.77-1.15)	0.534
AA	37 (10.1)	54 (14.8)	0.60 (0.38-0.97)	0.036	48 (9.3)	68 (13.6)	0.66 (0.44-0.99)	0.044	85 (9.6)	122 (14.1)	0.63 (0.46-0.86)	0.004
Dominant			0.79 (0.59-1.06)	0.111			0.92 (0.72-1.18)	0.517			0.86 (0.71-1.04)	0.112
Recessive			0.65 (0.42-1.02)	0.060			0.65 (0.44-0.97)	0.033			0.65 (0.49-0.88)	0.004
Co-dominant			0.80 (0.65-0.99)	0.039			0.87 (0.73-1.05)	0.140			0.84 (0.73-0.96)	0.012

aOR: Adjusted odds ratio; CI: confidence interval. Calculated by unconditional logistic analysis, adjusted for age and gender.

Table IV. *Docking protein 3 (DOK3) genotypes of cases and controls, and their associations with the risk of colorectal cancer according to the primary site.*

	Discovery				Validation				Combined			
DOK3 rs2279398 genotype	Cases, n (%)	Controls, n (%)	aOR (95% CI)	*p*-Value	Cases, n (%)	Controls, n (%)	aOR (95% CI)	*p*-Value	Cases, n (%)	Controls, n (%)	aOR (95% CI)	*p*-Value
Colonic cancer												
GG	108 (55.1)	158 (43.3)	1.00		147 (48.2)	237 (47.4)	1.00		255 (50.9)	395 (45.7)	1.00	
GA	69 (35.2)	153 (41.9)	0.70 (0.48-1.03)	0.070	130 (42.6)	195 (39.0)	1.08 (0.79-1.46)	0.631	199 (39.7)	348 (40.2)	0.90 (0.71-1.14)	0.389
AA	19 (9.7)	54 (14.8)	0.53 (0.30-0.96)	0.034	28 (9.2)	68 (13.6)	0.66 (0.41-1.07)	0.092	47 (9.4)	122 (14.1)	0.60 (0.41-0.87)	0.007
Dominant			0.66 (0.46-0.94)	0.021			0.97 (0.73-1.29)	0.826			0.82 (0.66-1.02)	0.082
Recessive			0.62 (0.36-1.09)	0.095			0.64 (0.40-1.01)	0.057			0.63 (0.44-0.89)	0.010
Co-dominant			0.70 (0.53-0.92)	0.009			0.89 (0.71-1.10)	0.269			0.80 (0.68-0.95)	0.009
Rectal cancer												
GG	74 (43.8)	158 (43.3)	1.00		107 (51.2)	237 (47.4)	1.00		181 (47.9)	395 (45.7)	1.00	
GA	77 (45.6)	153 (41.9)	1.07 (0.72-1.58)	0.751	82 (39.2)	195 (39.0)	0.93 (0.66-1.31)	0.677	159 (42.1)	348 (40.2)	0.99 (0.77-1.28)	0.945
AA	18 (4.9)	54 (14.8)	0.71 (0.39-1.30)	0.265	20 (9.6)	68 (13.6)	0.65 (0.37-1.12)	0.121	38 (10.0)	122 (14.1)	0.68 (0.45-1.01)	0.057
Dominant			0.97 (0.67-1.41)	0.883			0.86 (0.62-1.18)	0.349			0.91 (0.71-1.16)	0.442
Recessive			0.69 (0.39-1.21)	0.194			0.67 (0.40-1.13)	0.135			0.68 (0.46-0.99)	0.048
Co-dominant			0.99 (0.79-1.24)	0.954			1.09 (0.90-1.32)	0.393			1.05 (0.90-1.21)	0.553

aOR: Adjusted odds ratio; CI: confidence interval. Calculated by unconditional logistic analysis, adjusted for age and gender.

diseases also considerably increases the risk of CRC. More recently, based on the extensive study of genetic alterations in CRC, tumorigenesis has been identified as a multistep process that involves the accumulation of mutations in tumor-suppressor genes and oncogenes such as adenomatous polyposis coli (*APC*), *p53*, and *KRAS* (31).

The present study also showed that patients at high risk for developing CRC might be identified by genetic testing for specific polymorphisms. This finding can potentially benefit patients through the identification of the *DOK3* polymorphism, enabling the early detection of disease and providing information for appropriate treatment strategies. However, even though the present data identified certain gene variants as statistically significant risk factors, these results should be interpreted cautiously. First of all, when considering the multiple comparison issue, the possibility of a type I error cannot be ruled-out in the analysis of individual SNPs. Therefore, additional studies with larger sample sizes are required. Secondly, the present finding that the DOK3 polymorphism itself alters the protein expression still needs to be confirmed. In addition, gene interactions with environmental and lifestyle factors can also affect the relevance of different variants in cancer susceptibility. And last but not least, the current results do not conclusively reveal the exact role of the *DOK3* polymorphism in CRC tumorigenesis or its relationship with the RAS-ERK pathway.

The current findings indicate that genetic variations of *DOK3* may influence the risk of CRC. However, since the exact mechanism and function of these gene variants have not yet been fully defined, the present findings need to be confirmed in further studies with other populations in order to clarify the association between these polymorphisms and the risk of CRC.

Conflicts of Interest

The Authors declare no conflicts of interest with regard to this study

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Korean Ministry of Education, Science and Technology (NRF-2011-0015862 and NRF- 2014M3A9E2064642).

References

1 Brenner H, Kloor M and Pox CP: Colorectal cancer. Lancet *383*: 1490-1502, 2014.

2 Kraus S, Nabiochtchikov I, Shapira S and Arber N: Recent advances in personalized colorectal cancer research. Cancer Lett *347*: 15-21, 2014.

3 Jia WH, Zhang B, Matsuo K, Shin A, Xiang YB, Jee SH, Kim DH, Ren Z, Cai Q, Long J, Shi J, Wen W, Yang G, Delahanty RJ, Ji BT, Pan ZZ, Matsuda F, Gao YT, Oh JH, Ahn YO, Park EJ, Li HL, Park JW, Jo J, Jeong JY, Hosono S, Casey G, Peters U, Shu XO, Zeng YX and Zheng W: Genome-wide association analyses in East Asians identify new susceptibility loci for colorectal cancer. Nat Genet *45*: 191-196, 2013.

4 de la Chapelle A: Genetic predisposition to colorectal cancer. Nat Rev Cancer *4*: 769-780, 2004.

5 Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*: 281-297, 2004.

6 Bartel DP: MicroRNAs: target recognition and regulatory functions. Cell *136*: 215-233, 2009.

7 Calin GA and Croce CM: MicroRNA signatures in human cancers. Nat Rev Cancer *6*: 857-866, 2006.

8 Esquela-Kerscher A and Slack FJ: Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer *6*: 259-269, 2006.

9 Nakajima G, Hayashi K, Xi Y, Kudo K, Uchida K, Takasaki K, Yamamoto M and Ju J: Non-coding MicroRNAs hsa-let-7g and hsa-miR-181b are Associated with Chemoresponse to S-1 in Colon Cancer. Cancer Genomics Proteomics *3*: 317-324, 2006.

10 Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, Yuen ST, Chan TL, Kwong DL, Au GK, Liu CG, Calin GA, Croce CM and Harris CC: MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. JAMA *299*: 425-436, 2008.

11 Slaby O, Svoboda M, Fabian P, Smerdova T, Knoflickova D, Bednarikova M, Nenutil R and Vyzula R: Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer. Oncology *72*: 397-402, 2007.

12 Slaby O, Bienertova-Vasku J, Svoboda M and Vyzula R: Genetic polymorphisms and microRNAs: new direction in molecular epidemiology of solid cancer. J Cell Mol Med *16*: 8-21, 2012.

13 Chae YS, Kim JG, Lee SJ, Kang BW, Lee YJ, Park JY, Jeon HS, Park JS and Choi GS: A miR-146a polymorphism (rs2910164) predicts risk of and survival from colorectal cancer. Anticancer Res *33*: 3233-3239, 2013.

14 Lee HC, Kim JG, Chae YS, Sohn SK, Kang BW, Moon JH, Jeon SW, Lee MH, Lim KH, Park JY, Choi GS and Jun SH: Prognostic impact of microRNA-related gene polymorphisms on survival of patients with colorectal cancer. J Cancer Res Clin Oncol *136*: 1073-1078, 2010.

15 Lee SJ, Seo JW, Chae YS, Kim JG, Kang BW, Kim WW, Jung JH, Park HY, Jeong JY and Park JY: Genetic polymorphism of miR-196a as a prognostic biomarker for early breast cancer. Anticancer Res *34*: 2943-2949, 2014.

16 Krol J, Loedige I and Filipowicz W: The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet *11*: 597-610, 2010.

17 Mishra PJ, Humeniuk R, Longo-Sorbello GS, Banerjee D and Bertino JR: A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. Proc Natl Acad Sci USA *104*: 13513-13518, 2007.

18 Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS and Cheung VG: Genetic analysis of genome-wide variation in human gene expression. Nature 430: 743-747, 2004.

19 Dong LM, Potter JD, White E, Ulrich CM, Cardon LR and Peters U: Genetic susceptibility to cancer: the role of polymorphisms in candidate genes. JAMA *299*: 2423-2436, 2008.

20 Edge SB, Sobin LH, Page DL, Gospodarowicz MK, Greene FL and Winchester DP: Re: Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. J Natl Cancer Inst *97*: 463-4; author reply 4-5, 2005.

21 Wojcicka A, de la Chapelle A and Jazdzewski K: MicroRNA-related sequence variations in human cancers. Hum Genet *133*: 463-469, 2014.
22 Preskill C and Weidhaas JB: SNPs in microRNA binding sites as prognostic and predictive cancer biomarkers. Crit Rev Oncog *18*: 327-340, 2013.
23 Landi D, Gemignani F, Naccarati A, Pardini B, Vodicka P, Vodickova L, Novotny J, Forsti A, Hemminki K, Canzian F and Landi S: Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. Carcinogenesis *29*: 579-584, 2008.
24 Landi D, Moreno V, Guino E, Vodicka P, Pardini B, Naccarati A, Canzian F, Barale R, Gemignani F and Landi S: Polymorphisms affecting micro-RNA regulation and associated with the risk of dietary-related cancers: a review from the literature and new evidence for a functional role of rs17281995 (CD86) and rs1051690 (INSR), previously associated with colorectal cancer. Mutat Res *717*: 109-115, 2011.
25 Naccarati A, Pardini B, Stefano L, Landi D, Slyskova J, Novotny J, Levy M, Polakova V, Lipska L and Vodicka P: Polymorphisms in miRNA-binding sites of nucleotide excision repair genes and colorectal cancer risk. Carcinogenesis *33*: 1346-1351, 2012.
26 Lemay S, Davidson D, Latour S and Veillette A: Dok-3, a novel adapter molecule involved in the negative regulation of immunoreceptor signaling. Mol Cell Biol *20*: 2743-2754, 2000.
27 Spitzner M, Emons G, Kramer F, Gaedcke J, Rave-Frank M, Scharf JG, Burfeind P, Becker H, Beissbarth T, Ghadimi BM, Ried T and Grade M: A gene expression signature for chemoradiosensitivity of colorectal cancer cells. Int J Radiat Oncol Biol Phys *78*: 1184-1192, 2010.
28 Van Cutsem E, Kohne CH, Lang I, Folprecht G, Nowacki MP, Cascinu S, Shchepotin I, Maurel J, Cunningham D, Tejpar S, Schlichting M, Zubel A, Celik I, Rougier P and Ciardiello F: Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status. J Clin Oncol *29*: 2011-2019, 2011.
29 Zhang W, Winder T, Ning Y, Pohl A, Yang D, Kahn M, Lurje G, Labonte MJ, Wilson PM, Gordon MA, Hu-Lieskovan S, Mauro DJ, Langer C, Rowinsky EK and Lenz HJ: A let-7 microRNA-binding site polymorphism in 3'-untranslated region of KRAS gene predicts response in wild-type KRAS patients with metastatic colorectal cancer treated with cetuximab monotherapy. Ann Oncol *22*: 104-109, 2011.
30 Tarraga Lopez PJ, Albero JS and Rodriguez-Montes JA: Primary and secondary prevention of colorectal cancer. Clin Med Insights Gastroenterol *7*: 33-46, 2014.
31 Zoratto F, Rossi L, Verrico M, Papa A, Basso E, Zullo A, Tomao L, Romiti A, Lo Russo G and Tomao S: Focus on genetic and epigenetic events of colorectal cancer pathogenesis: implications for molecular diagnosis. Tumour Biol *35*: 6195-6206, 2014.

Received April 12, 2016
Revised May 17, 2016
Accepted May 24, 2016