Original Article

# Identification of Candidate Genes Carrying Polymorphisms Associated With the Risk of Colorectal Cancer by Analyzing the Colorectal Mutome and MicroRNAome

Debora Landi, PhD[1]; Federica Gemignani, PhD[1]; Barbara Pardini, PhD[2]; Alessio Naccarati, PhD[2]; Sonia Garritano, PhD[1]; Pavel Vodicka, MD, PhD[2,3]; Ludmila Vodickova, PhD[2,4]; Federico Canzian, PhD[5]; Jan Novotny, MD[6,7]; Roberto Barale, PhD[1]; and Stefano Landi, PhD[1]

**BACKGROUND:** The presence of single-nucleotide polymorphisms (SNPs) within the 3′-untranslated regions of genes could affect the binding between a microRNA (miRNA) and its target, with consequences on gene expression regulation. Considering the important role of miRNAs in carcinogenesis, it is hypothesized here that these SNPs could also affect the individual risk of colorectal cancer (CRC). **METHODS:** To test this hypothesis, a list was developed of 140 somatically mutated genes deduced from previous works on the mutome of the CRC. A further selection was conducted of SNPs within target sites for miRNAs that are expressed only in the colorectum (the colorectal microRNAome) and having adequate population frequencies. This yielded 12 SNPs that were genotyped in a case-control association study on 717 colorectal cases and 1171 controls from the Czech Republic. **RESULTS:** Statistically significant associations were found between the risk of CRC and the variant alleles of *KIAA0182* (rs709805) (odds ratio = 1.57; 95% confidence interval = 1.06-2.78, for the variant homozygotes) and *NUP210* genes (rs354476) (odds ratio = 1.36; 95% confidence interval = 1.02-1.82, for the variant homozygotes). **CONCLUSIONS:** The results support the study hypothesis and highlight the importance of SNPs within miRNA-dependent regulatory regions. Further studies on the role exerted by *NUP210* and *KIAA0182* in colorectal carcinogenesis are warranted. ***Cancer*** **2012;118:4670-80.** © *2012 American Cancer Society.*

**KEYWORDS:** colorectal cancer, single-nucleotide polymorphism, microRNA, KIAA0182, NUP210.

Colorectal cancer (CRC) is the third most common cause of cancer death in the world. There is a consistent body of evidence from prospective studies to indicate that a high glucidic diet,[1] high consumption of fat and red meat,[2] and obesity are positively associated with CRC risk.[3] Conversely, the use of nonsteroidal anti-inflammatory drugs[4] and the consumption of brassica vegetables[5] significantly reduce the risk for this cancer. Genetic factors can also contribute in modulating risk of CRC. In fact, high-risk allelic variants are responsible for familial cases (such as in familial amyloid polyneuropathy or hereditary nonpolyposis colorectal cancer), whereas in the sporadic forms, the risk can be modulated by low-risk variants (genetic polymorphisms). Recently, several genome-wide association studies were performed and risk alleles were identified,[6-9] such as those reviewed by Houlston et al.[10] The genome-wide association studies offer the advantage of exploring thousands of loci without the need to establish any a priori hypothesis. However, they also suffer from some limitations, including the need for adjustment of the *P* values for multiple testing (lowering the power of the study) and the difficulty of interpreting the results or generating any hypothesis on anonymous markers associated with risk.[11] In this context, the studies carried out on candidate genes under the classical hypothesis "common allele, common disease" may provide advantages.[12] Actually, these studies helped detect the null genotype of the gene *GSTM1* (glutathione S-transferase mu 1),[13] and specific single-nucleotide polymorphisms (SNPs) within, eg, *MTHFR* (methylene tetrahydrofolate reductase), *NOD2* (nucleotide-binding oligomerization domain 2),[14] or *TP53* (tumor protein p53)[15] as risk factors for CRC.

Gene deregulation is one of the key mechanisms by which cells can progress to cancer. The post-transcriptional regulation carried out by microRNAs (miRNAs) is one of the most interesting and powerful of these mechanisms.[16] SNPs can reside within the genes encoding for miRNAs or within the target sequences at the 3′-untranslated regions (3′-UTRs); thus, they can affect the strength of the binding between an miRNA and its target.[17,18] As a result, SNPs can affect gene regulation, and therefore it is conceivable that these SNPs could be associated with a differential risk of cancer, as shown in a previous study.[19] Following this concept, we report here a case-control association study on CRC where SNPs were

**Corresponding author:** Stefano Landi, PhD, Department of Biology, University of Pisa, Via Derna 1, 56100 Pisa, Italy fax: (39) 050 2211527; slandi@biologia.unipi.it

[1]Department of Biology, University of Pisa, Pisa, Italy; [2]Department of Molecular Biology of Cancer, Institute of Experimental Medicine, Prague, Czech Republic; [3]Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, Czech Republic; [4]Centre of Occupational Health, National Institute of Public Health, Prague, Czech Republic; [5]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany; [6]Department of Oncology, First Faculty of Medicine, Charles University, Prague, Czech Republic; [7]Oncology Centre, Hospital Příbram, Příbram, Czech Republic

**DOI:** 10.1002/cncr.27435, **Received:** May 13, 2011; **Revised:** October 31, 2011; **Accepted:** November 18, 2011, **Published online** January 26, 2012 in Wiley Online Library (wileyonlinelibrary.com)

selected on the basis of their potential effect on the different binding between miRNAs and their targets. In order to restrict the analysis to SNPs with a high likelihood of being associated with the risk of CRC, we reviewed the literature and filtered the results using a series of stringent a priori hypotheses.

Initially, we started from a list of 140 somatically mutated genes thought to be crucial in driving the development of CRC (Table 1). These genes derive from the analysis of the mutome, where 20,857 transcripts from 18,191 genes were sequenced in 11 patients with CRC.[20] Then, we predicted in silico the target sequences for the miRNAs within the 3′-UTRs and indexed all the SNPs falling within these targets. Finally, we considered only those polymorphisms affecting the binding with miRNAs expressed specifically in the colon and rectum (Table 1), according to the colorectal microRNAome from Cummins et al.[21] Thus, the overlap between the mutome and the microRNAome data sets allowed the selection of only 12 SNPs. These SNPs were checked for their association with the risk of CRC in a case-control study on 717 CRC cases and 1171 controls from the Czech Republic.

## MATERIALS AND METHODS

### Study Population

Cases were patients with histologically confirmed CRC recruited between September 2004 and February 2009 from 9 oncological departments in the Czech Republic: Prague (2 departments), Benešov, Brno, Liberec, Ples, Příbram, Ústí nad Labem, and Zlín. During this period, a total of 968 cases provided blood samples. This study includes 717 subjects (74% of the whole set) who were able to be interviewed, provided biological samples, and who were genotyped appropriately. The mean age at diagnosis of the patients was 61.9 years.

Controls were 739 hospital-based volunteers with negative colonoscopy results for malignancy or idiopathic bowel diseases (cancer-free colonoscopy inspected controls [CFCCs]). CFCCs were selected among individuals admitted to the same hospitals during the same period of recruitment of the cases. The reasons for undergoing the colonoscopy were: 1) positive fecal occult blood test, 2) hemorrhoids, 3) abdominal pain of unknown origin, and 4) macroscopic bleeding. Cases and CFCCs had the same inclusion and exclusion criteria. Among 739 CFCCs, 502 (83.1%) showed complete covariates and valid genotypes and were analyzed in this study. The mean age at the time of sampling was 55.8 years.

A second group of controls consisted of 669 healthy blood donor volunteers (HBDV) collected from a blood donor center in Prague. All individuals were subjected to standard examinations to verify the health status for blood donation (detailed blood count, urinary examination, blood pressure, and general examination). The sample collection was performed at the same time as that of the other 2 study groups above. The mean age at the time of sampling was 49.2 years. All subjects were informed and provided written consent to participate in the study and to approve the use of their biological samples for genetic analyses, according to the Helsinki declaration. The design of the study was approved by the local ethics committee. Cases and controls were personally 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, diabetes, and family/personal history of cancer. A portion of the cases and controls presented here were also analyzed in previous association studies.[22,23]

### Selection of Candidate Genes

Wood et al carried out a mutome study on 11 patients with CRC.[20] In order to discriminate between passenger and driver mutations, each mutation was further verified in an independent series of 96 patients with CRC. Wood et al[20] reported a list of 140 candidate genes ("*CAN-genes*" in Table 1) for driving carcinogenesis in the colorectum. The initial selection of the present study was based on this list.

### Selection of MicroRNA Targets

We selected the 3′-UTRs of the 140 "CAN-genes", according to the University of California Santa Cruz genome browser (http://genome.ucsc.edu). Then, the putative miRNA binding sites were identified by means of specialized algorithms: miRBase (http://www.mirbase.org/),[24] miRanda (http://www.microrna.org/),[25] MicroInspector (http://mirna.imbb.forth.gr/microinspector/),[26] Diana-Micro-T (http://www.diana.pcbi.upenn.edu/cgi-bin/micro_t.cgi),[27] and TargetScan Human 5.1 (http://www.targetscan.org/),[28] using the default parameters included in the software.

### Selection of SNPs

The predicted miRNA binding sites were screened for the presence of SNPs by an extensive search in the SNP database (dbSNP; http://www.ncbi.nlm.nih.gov/SNP/). As a result, we found 61 SNPs within 31 genes. In order to have an appropriate statistical power, we excluded the SNPs having the minor allele frequency lower than 0.24 in Caucasians, and 37 SNPs were retained (Table 2).

**Table 1.** List of 140 Candidate (CAN)-Genes for Colorectal Cancer[20] and 190 MicroRNAs Expressed in the Colorectum[21] From Which the Single-Nucleotide Polymorphism Selection Was Based

### CAN-Genes Selected (Colorectal Mutome)

| | | | | | |
|---|---|---|---|---|---|
| *ABCA1* | *CD109* | *GALNS* | *LGR6* | *PLCG2* | *TAF2* |
| *ABCB11* | *CHL1* | *GLI3* | *LMO7* | *PRDM9* | *TBX22* |
| *ACSL5* | *CHR415SYT* | *GNAS* | *MAP1B* | *PRKD1* | *TCERG1L* |
| *ADAM19* | *CLSTN2* | *GPR112* | *MAP2* | *PTEN* | *TCF7L2* |
| *ADAM29* | *CNTN4* | *GPR158* | *MAP2K7* | *PTPRD* | *TGFBR2* |
| *ADAMTS18* | *COL3A1* | *GRID1* | *MAPK8IP2* | *PTPRS* | *TGM3* |
| *ADAMTS20* | *CPAMD8* | *GRM1* | *MCM3AP* | *PTPRU* | *TIAM1* |
| *ADAMTSL3* | *CSMD3* | *GUCY1A2* | *MGC20470* | *RASGRF2* | *TLR9* |
| *ADARB2* | *CUTL1* | *HAPIP* | *MKRN3* | *RET* | *TNN* |
| *AGC1* | *CX40.1* | *HAPLN1* | *MMP2* | *ROBO1* | *TP53* |
| *AKAP12* | *DPP10* | *HIST1H1B* | *MYO18B* | *RUNX1T1* | *UHRF2* |
| *AKAP6* | *DSCAML1* | *IGFBP3* | *MYO5C* | *SCN3B* | *UQCRC2* |
| *ALK* | *DTNB* | *IGSF22* | *MYOHD1* | *SEC8L1* | *ZNF262* |
| *APC* | *EDD1* | *IRS4* | *NAV3* | *SEMA3D* | *ZNF442* |
| *ARHGEF10* | *EPHA3* | *ITGAE* | *NF1* | *SFRS6* | *ZNF521* |
| *ATP11A* | *EPHB6* | *K6IRS3* | *NOS3* | *SH3TC1* | |
| *ATP13A1* | *ERCC6* | *KCNQ5* | *NTNG1* | *SHANK1* | |
| *ATP13A5* | *EYA4* | *KIAA0182* | *NUP210* | *SLC22A15* | |
| *BCL9* | *F8* | *KIAA0367* | *OR51E1* | *SLC29A1* | |
| *C10orf137* | *FBN2* | *KIAA0556* | *P2RX7* | *SMAD2* | |
| *C13orf7* | *FBXW7* | *KIAA1409* | *PCDH11X* | *SMAD3* | |
| *C14orf115* | *FLJ10404* | *KIAA2022* | *PCDHA9* | *SMAD4* | |
| *C15orf2* | *FLJ13305* | *KRAS* | *PIK3CA* | *SMTN* | |
| *C1QR1* | *FLNC* | *LAMA1* | *PKNOX1* | *SORL1* | |
| *CACNA2D3* | *FN1* | *LCN9* | *PLB1* | *STAB1* | |

### miRNA Expressed in Colorectum (the Colorectal MicroRNAome)

| | | | | | |
|---|---|---|---|---|---|
| MIR-21 | MIR-27B | MIR-218 | LET-7C | MIR-18B | MIR-302C |
| MIR-200A | LET-7F | MIR-130A | MIR-193B | MIR-302B | MIR-302D |
| MIR-200C | MIR-374 | MIR-93 | MIR-223 | MIR-199B | MIR-135A |
| MIR-143 | MIR-18A | MIR-182 | MIR-186 | MIR-499 | MIR-30A-3P |
| MIR-200B | LET-7A | MIR-210 | MIR-342 | MIR-409-3P | MIR-20B |
| MIR-27A | MIR-335 | MIR-150 | MIR-199A | MIR-425 | MIR-495 |
| MIR-24 | MIR-92 | MIR-15B | MIR-181B | MIR-153 | MIR-148B |
| MIR-26A | MIR-106B | LET-7I | MIR-214 | MIR-205 | MIR-184 |
| MIR-145 | MIR-15A | MIR-320 | MIR-99B | MIR-219 | MIR-204 |
| MIR-101 | MIR-33 | MIR-424 | MIR-148A | MIR-324-5P | MIR-299-5P |
| MIR-194 | MIR-372 | MIR-301 | MIR-497 | MIR-339 | MIR-326 |
| MIR-192 | MIR-23B | MIR-142-5P | MIR-331 | MIR-138 | MIR-34C |
| MIR-451 | MIR-373 | MIR-429 | MIR-7 | MIR-181D | MIR-379 |
| MIR-22 | MIR-30B | MIR-96 | MIR-34A | MIR-410 | MIR-382 |
| MIR-103 | MIR-126 | MIR-151 | MIR-155 | MIR-422B | MIR-449 |
| MIR-16 | MIR-19A | MIR-375 | MIR-128B | MIR-190 | MIR-512-1 5P |
| MIR-25 | MIR-30C | MIR-9 | MIR-128A | MIR-378 | MIR-512-2 5P |
| MIR-141 | MIR-26B | MIR-371 | MIR-135B | MIR-367 | MIR-494 |
| MIR-29B | MIR-32 | MIR-215 | MIR-152 | MIR-146B | MIR-498 |
| MIR-30E-5P | MIR-125B | MIR-100 | MIR-98 | MIR-132 | MIR-502 |
| MIR-142-3P | MIR-221 | MIR-133A | MIR-450 | MIR-377 | MIR-507 |
| MIR-191 | MIR-107 | MIR-28 | LET-7E | MIR-124A | MIR-517A |
| MIR-193A | MIR-29C | MIR-30E-3P | MIR-125A | MIR-299-3P | MIR-517C |
| MIR-144 | MIR-10A | MIR-203 | MIR-136 | MIR-452 | MIR-518B |
| MIR-222 | MIR-181A | LET-7D | MIR-183 | MIR-501 | MIR-520D |
| MIR-19B | MIR-17-5P | MIR-95 | MIR-197 | MIR-514 | MIR-520G |
| MIR-30D | LET-7B | MIR-140 | MIR-363 | MIR-213 | MIR-520H |
| MIR-31 | MIR-20A | MIR-196B | MIR-139 | MIR-345 | MIR-521 |
| MIR-30A-5P | MIR-17-3P | MIR-338 | MIR-106A | MIR-381 | MIR-525 |
| MIR-23A | MIR-195 | MIR-365 | MIR-185 | MIR-224 | MIR-525 |
| LET-7G | MIR-1 | MIR-149 | MIR-181C | MIR-296 | |
| MIR-29A | MIR-130B | MIR-330 | MIR-188 | MIR-302A | |

**Table 2.** Selected Single-Nucleotide Polymorphisms (SNPs) with Minor Allele Frequency (MAF) > 0.25 and SNPs in MicroRNA Targets

| Gene ID | SNPs in the 3′-UTR (Generic) | Variation | MAF (Caucasian) | miRNA From the MicroRNAome[21] Predicted to Bind the Target | ΔGwt (kcal/mol) | ΔGvar (kcal/mol) | ΔΔG | \|ΔΔG\| | $\|\Delta\Delta G\|_{Tot}$ |
|---|---|---|---|---|---|---|---|---|---|
| *ABCA1* | rs4149338 | C/T | 0.28 | none | | | | | |
| | rs4149339 | C/T | 0.42 | none | | | | | |
| *ABCB11* | rs496550 | G/A | 0.40 | miR-363 (microInspector) | −13,18 | −12.35 | +0.83 | 0.83 | 0.83 |
| | rs473351 | G/A | 0.43 | miR-182 (microInspector) | −18,17 | −14.45 | +3.72 | 3.72 | 3.72 |
| | rs495714 | A/G | 0.40 | miR-324-3p (TargetScan Human 5.1) | −10,58 | −15.77 | −5.19 | 5.19 | 6.07 |
| | | | | miR-196b (microcosm; microInspector) | −19,60 | −19.16 | +0.44 | 0.44 | |
| | | | | miR-196a (micoInspector) | −16,60 | −16.16 | +0.44 | 0.44 | |
| *ACSL5* | rs8624 | C/T | 0.30 | none | | | | | |
| *ADAMTSL3* | rs2135551 | A/G | 0.29 | none | | | | | |
| *ADARB2* | rs1046914 | A/G | 0.35 | none | | | | | |
| | rs904960 | C/T | 0.35 | miR-32 (microcosm; microInspector) | −15.20 | −15.72 | −0.52 | 0.52 | 2.34 |
| | | | | miR-25 (microInspector) | −17.89 | −17.01 | +0.88 | 0.88 | |
| | | | | miR-367 (microInspector) | −14.15 | −13.74 | +0.41 | 0.41 | |
| | | | | miR-363 (microInspector) | −14.04 | −14.57 | −0.53 | 0.53 | |
| | rs1129227 | T/C | 0.32 | none | | | | | |
| *ATP11A* | rs3742232 | T/C | 0.41 | none | | | | | |
| *C1QR1* | rs2749813 | G/A | 0.36 | none | | | | | |
| | rs6076019 | T/C | 0.45 | none | | | | | |
| *CD109* | rs3012518 | G/A | 0.44 | miR-299-3p (TargetScan Human 5.1) | −18.94 | −17.74 | +1.20 | 1.20 | 1.20 |
| *CSMD3* | rs5894075 | G/T | 0.32 | none | | | | | |
| *DSCAML1* | rs2925768 | G/A | 0.38 | none | | | | | |
| *EPHA3* | rs7650466 | C/T | 0.29 | none | | | | | |
| *EYA4* | rs3734279 | C/T | 0.38 | miR-203 (TargetScan Human 5.1) | −8.80 | −8.64 | +0.16 | 0.16 | 0.16 |
| *GALNS* | rs1141390 | T/C | 0.25 | none | | | | | |
| *IGSF22* | rs2289965 | A/G | 0.25 | miR-142-3p (TargetScan Human 5.1) | −14.63 | −14.24 | +0.39 | 0.39 | 1.65 |
| | | | | miR-324-5p (microInspector) | −19.70 | −18.44 | +1.26 | 1.26 | |
| *KIAA0182* | rs8571 | G/A | 0.37 | none | | | | | |
| | rs709805 | G/A | 0.35 | miR-324-3p (microInspector) | −25.78 | −24.07 | +1.71 | 1.71 | 1.71 |
| *KRAS* | rs712 | T/G | 0.39 | miR-200b (TargetScan Human 5.1) | −10.19 | −10.51 | −0.32 | 0.32 | 1.13 |
| | | | | miR-429 (TargetScan Human 5.1) | −9.20 | −9.58 | −0.38 | 0.38 | |
| | | | | miR-200c (TargetScan Human 5.1) | −11.28 | −11.45 | −0.17 | 0.17 | |
| | | | | miR-193b (TargetScan Human 5.1) | −9.83 | −10.09 | −0.26 | 0.26 | |
| | rs12245 | T/A | 0.44 | none | | | | | |
| *MAP2K7* | rs3679 | C/T | 0.35 | none | | | | | |
| *MMP2* | rs7201 | A/C | 0.50 | none | | | | | |
| *NUP210* | rs1048650 | C/T | 0.47 | miR-22 (miRAnda) | −19.97 | −20.57 | −0.60 | 0.60 | 0.60 |
| | rs354476 | T/C | 0.48 | miR-125a (DianaMicroT) | −13.27 | −13.32 | −0.05 | 0.05 | 0.86 |
| | | | | miR-125b (DianaMicroT; microInspector) | −10.84 | −10.03 | +0.81 | 081 | |
| *PKHD1* | rs2784198 | C/A | 0,38 | none | | | | | |
| *PKNOX1* | rs2839629 | G/A | 0,45 | miR-18a (miRAnda) | −15.46 | −15.18 | +0.28 | 0.28 | 0.67 |
| | | | | miR-18b (miRAnda) | −14.87 | −14.48 | +0.39 | 0.39 | |
| | rs378528 | C/T | 0,28 | none | | | | | |
| | rs2839628 | C/G | 0,26 | none | | | | | |
| | rs2839628 | C/G | 0,26 | none | | | | | |
| *SMAD3* | rs2278670 | C/T | 0,28 | none | | | | | |
| | rs3743343 | T/C | 0,27 | none | | | | | |
| *TGM3* | rs214832 | C/T | 0,43 | none | | | | | |
| | rs214831 | G/A | 0,36 | none | | | | | |

For each SNP in miRNA targets, the ΔG was calculated (COFOLD software, Vienna Package).
miRNA indicates microRNA; UTR, untranslated region.

As a third criterion of selection, we kept only those SNPs within target sites for miRNAs specifically expressed in CRC. These miRNAs were taken from a study on the microRNAome of the colorectum by Cummins et al[21] (the list is shown in Table 1). Thus, only 12 SNPs were selected and verified in the case-control association study (Table 2).

### SNP Genotyping

Genomic DNA was isolated from peripheral blood lymphocytes, using standard procedures. The DNA samples from cases and controls were randomly placed on plates where an equal number of cases and controls could be run simultaneously. Genotyping of the 12 selected SNPs was carried out by using the KASPar chemistry (KBioscience, Hoddesdon, UK), which is a competitive allele-specific polymerase chain reaction (PCR) SNP genotyping system that uses fluorescence resonance energy transfer quencher cassette oligonucleotides. The reaction employed the KASP 2× Reaction Mix, KASPar primers and probes, water, and 5 ng of DNA for 10 μL of reaction and a standard PCR protocol available from KBioscience. Duplicate samples (5%), no-template controls in each plate, and Hardy-Weinberg equilibrium tests were used as quality control tests.

### Bioinformatics

For the selected SNPs, the algorithm *RNAcofold* (http://rna.tbi.univie.ac.at/cgi-bin/RNAcofold.cgi) was run to assess the Gibbs binding free energy (ΔG, expressed in kilojoules per mole), both for the common and the variant alleles. The algorithm *RNAcofold* computes the hybridization energy and base-pairing pattern of 2 RNA sequences.[29] The difference of the free energies between the 2 alleles was computed as "variation of ΔG" (ie, ΔΔG) (Table 2). Because the neighbor sequence of each SNP can be a target for different miRNAs, we calculated the sum of the absolute values of ΔΔGs for each SNP (ie, $|\Delta\Delta G|_{tot} = \Sigma\ |\Delta\Delta G|$) (Table 2). The $|\Delta\Delta G|_{tot}$ should be considered as a sort of "disturbance index" predicting the likelihood for a given SNP to affect the function of the 3′-UTR, and it allows a ranking of SNPs for their relevance, as illustrated in previous studies.[30,31]

### Vectors Employed and In Vitro Assays

We PCR-amplified both the common and variant 3′-UTR regions of the *NUP210* and *KIAA0182* genes. The amplification was done using 2 primers having a sequence of 6 bases to their 5′ ends, recognized by the restriction enzyme SacI → gagctc (forward primer) and XhoI → ctcgag (reverse primer). The PCR products were cloned in the pUC57 vectors. Successively, the plasmids were cleaved with SacI and XhoI, and the inserts were cloned downstream from a reporter vector containing the firefly luciferase (*Photinus pyralis*) and the *Renilla* luciferase (*Renilla reniformis*) genes (pmiR-GLO vector; Promega, Madison, Wis). Caco2, HCT116_p53 WT, and HCT116_p53−/− cell lines were plated at a density of approximately $2 \times 10^5$ cells per well in 6-well plates and incubated overnight at 5% $CO_2$, 37°C in a humidified incubator. They were transiently transfected in 12 μL of PolyFect transfection reagent (2 mg/mL; Qiagen Spa, Italy) and 1.5 μg of luciferase/*Renilla* chimeric construct, according to the manufacturer's protocols. Each experimental point was repeated 6 times, and the experiment was repeated 3 times. Forty-eight hours after transfection, cells were lysed with 500 μL of 1× passive lysis buffer (dual-luciferase reporter assay kit; Promega, USA) after washing with phosphate-buffered saline. Cells were lysed for 15 minutes at room temperature, transferred to 1.5-mL microcentrifuge tubes, vortexed briefly, and centrifuged at 13,000 rpm for 30 seconds to pellet cell debris. Supernatants were transferred to clean tubes and used for the measure of activity of firefly and *Renilla* luciferases, using a dual-luciferase reporter assay kit and a luminometer (Berthold Technologies, Germany). The firefly luciferase (Luc) reporter was measured first by adding Luciferase Assay Reagent II (LARII). After the measurement of the firefly luminescence, this reaction was quenched, and the *Renilla* luciferase (Ren) reaction was simultaneously initiated by adding Stop & Glo Reagent (dual-luciferase reporter assay kit; Promega). The measurements of the luminescence of luciferase and *Renilla* of the nontransfected cells (background) were subtracted to the values obtained for the transfected cells with the pmiR-GLO vector containing the 3′-UTR. The luminescence of the *Renilla* luciferase was used as the reference value to calculate the value of firefly luciferase (Luc/Ren ratio of luminescence).

This ratio (Luc/Ren) was compared to the one obtained for the transfection with the pmiR-GLO vector without 3′-UTR (empty vector (EV)).

### Statistical Analyses

To verify whether the genotypes were in Hardy-Weinberg equilibrium in controls, we used the chi-square test (1 degree of freedom), with a type-I alpha error of 0.05. The multivariate logistic regression analysis (MLR) was used to test the association between genotypes and risk of CRC. The covariates included in the model were: sex, age, smoking habit (nonsmokers vs smokers and ex-smokers), body mass index, any positive familial history of CRC, education level (high, intermediate, and low), and living area (country, town neighborhood, and town). The individual SNPs were input in the MLR analysis; however, 3

**Table 3.** Haplotype Analyses of *ABCB11* and *NUP210* Loci

| Locus | Haplotype (Code) | Frequency (All Samples) | SNP1 | SNP2 | SNP3 |
|---|---|---|---|---|---|
| *ABCB11* | | | **rs496550** | **rs495714** | **rs473351** |
| | 1 | 0.526 | G | A | G |
| | 2 | 0.461 | A | G | A |
| | 3 | 0.012 | A | G | G |
| | 4 | 0.001 | G | G | G |
| **Linkage Disequilibrium** | | | vs SNP2: D′=1; $r^2$=0.995<br>vs SNP3 D′=1; $r^2$=0.954 | vs SNP3: D′=1; $r^2$=0.949 | |
| *NUP210* | | | **rs354476** | **rs1048650** | |
| | 1 | 0.526 | T | C | |
| | 2 | 0.436 | C | T | |
| | 3 | 0.038 | C | C | |
| **Linkage Disequilibrium** | | | vs SNP2: D′=1; $r^2$=0.86 | | |

SNP indicates single-nucleotide polymorphism.

and 2 SNPs were genotyped for *ABCB11* and *NUP210*, respectively. Thus, we first reconstructed the individual haplotypes for these 2 genes with the software Fastphase,[31] then we calculated the linkage disequilibrium (LD) between SNPs and found that they were strongly associated with each other ($r^2 > 0.85$). Thus, we used the SNPs that showed the highest values of $|\Delta\Delta G|_{tot}$ (ie, rs354476 for *NUP210* and rs495714 for *ABCB11*) because the others were almost completely tagged by them. The association between SNPs and CRC risk was calculated, by estimating the odds ratio (OR) and its 95% confidence interval (CI), adjusted for both continuous and discontinuous covariates, as linear variables (the adjusted OR). For all genotypes, we performed the Cochran-Armitage trend test[32] in order to detect the best genetic model (dominant, additive, recessive), and the one with the highest likelihood was input in the MLR analysis. The statistical threshold of significance was set at 0.05; however, the more restrictive Bonferroni correction was applied. In this case, because of the strong LD between SNPs, only 9 completely independent statistics were performed on the genotypes, and this value was used for the Bonferroni correction ($P$ threshold $= 5.56 \times 10^{-3}$). For the in vitro assays, the ratio of the measurements of fluorescence (Luc/Ren), subtracted of the background, were compared among cell lines and between genotypes (for each gene) using the multifactor analysis of variance (MANOVA). All statistical tests were 2-tailed and were carried out using Statgraphics Centurion software (StatPoint Technologies, Warrenton, Va).

## RESULTS

### *Study Group and Genotype Analysis*

Table 1 shows the initial list of genes, whereas Table 2 shows the calculations of the selected SNPs for their $|\Delta\Delta G|_{tot}$. The definitive number of subjects for whom all the data were available accounted for 717 cases and 1171 controls. Among the controls, 502 were CFCCs and 669 were HBDVs. The quality control of genotypes was assured (>99% concordance) and all the SNPs were in Hardy-Weinberg equilibrium (data not shown). The SNPs within *ABCB11* and *NUP210* were analyzed for their LD, and the haplotypes are reported in Table 3. The SNPs within *ABCB11* and within *NUP210* show a strong LD to each other ($r^2 > 0.85$), with the prevalence of 2 main haplotypes for each gene. Thus, for further analyses, as "tagging SNP" for each gene, we used the SNP showing the highest $|\Delta\Delta G|_{tot}$ (Table 2).

The characteristics of the study population as well as the outcomes from MLR analyses are given in Table 4. When the CFCCs were used as reference group, the risk of CRC was associated in a statistically significant way with an increased age and a positive history of smoking habit: the cases were, on average, approximately 6 years older than CFCCs, whereas the never-smokers represented 51.2% of the cases and 58.8% of the controls. Two genotypes were also associated with CRC risk: the AA homozygotes for rs709805 (*KIAA0182*) showed an OR of 1.72 (95% CI = 1.06-2.78; $P = 2.8 \times 10^{-2}$), as compared to the GG+GA group, and the CC homozygotes for rs354476 (*NUP210*) had an OR of 1.36 (95% CI = 1.02-1.82; $P = 4.5 \times 10^{-3}$), compared with the TT+TC group. This latter SNP was also significantly associated with the risk of CRC after applying the Bonferroni correction. The fact that only the homozygotes were the genotypes at risk prevented us from finding a statistically significant difference between alleles, when the analyses were carried out on per-allele bases (rs709805, minor allele frequency = 0.29 and 0.26; rs354476, 0.48 and 0.45 among cases and controls, respectively). When the

**Table 4.** Characteristics of Colorectal Cancer Patients and Control Subjects and Adjusted Odds Ratio (OR) and 95% Confidence Interval (CI) Following Multivariate Logistic Regression Analysis

| Characteristic | Cases n (%) | CFCCs n (%) | Adjusted OR (95% CI) | *P* | CFCC+HBDV n (%) | Adjusted OR (95% CI) | *P* |
|---|---|---|---|---|---|---|---|
| **Sex** | | | | | | | |
| Male | 429 (59.8) | 274 (54.6) | Ref | | 645 (55.1) | Ref | |
| Female | 288 (40.2) | 228 (45.4) | 0.91 (0.71-1.18) | .48 | 526 (44..9) | 1.08 (0.86-1.36) | 0.51 |
| **Age (average)** | 61.9 | **55.8** | **1.05 (1.04 -1.06)** | $<10^{-6}$ | **52.0** | **1.09 (1.08-1.11)** | $<10^{-6}$ |
| ≤45 years old | 56 (7.8) | 110 (21.9) | | | 429 (36.6) | | |
| >45 years old | 765 (92.2) | 392 (78.1) | | | 742 (63.4) | | |
| **Smoking habit** | | | | | | | |
| Nonsmokers | 367 (51.2) | 295 (58.8) | Ref | | 686 (58.6) | Ref | |
| Smokers + Ex-smoker | 350 (48.8) | **207 (41.2)** | **1.41 (1.10-1.82)** | $7.8\times10^{-3}$ | **485 (41.4)** | **1.48 (1.18-1.84)** | $5.4\times10^{-4}$ |
| **BMI (average)** | 26.3 | 26.0 | 0.98 (0.96-1.02) | 0.58 | 26.2 | 0.99 (0.97-1.02) | 0.43 |
| BMI ≤ median (26.2) | 410 (57.2) | 255 (50.8) | | | 637 (54.4) | | |
| BMI > median (26.2) | 307 (42.8) | 247 (49.2) | | | 534 (45.6) | | |
| **Positive familial history of CRC** | | | | | | | |
| Yes | 117 (16.3) | 77 (15.3) | Ref | | 119 (10.2) | Ref | |
| No | 600 (83.7) | 425 (84.7) | 0.85 (0.61-1.18) | 0.32 | **1052 (89.8)** | **0.56 (0.41-0.76)** | $2.3\times10^{-4}$ |
| **Education** | | | | | | | |
| High | 112 (15.6) | 94 (18.7) | Ref | | 292 (24..9) | Ref | |
| Intermediate | 385 (53.7) | 272 (54.2) | 1.32 (0.94-1.87) | 0.12 | **697 (59.6)** | **1.50 (1.12-2.01)** | $6.6\times10^{-3}$ |
| Low | 220 (30.7) | 136 (27.1) | 1.23 (0.84-1.81) | 0.29 | **182 (15.5)** | **1.74 (1.23-2.45)** | $1.6\times10^{-3}$ |
| **Living area** | | | | | | | |
| Country | 193 (26.9) | 123 (24.5) | Ref | | 219 (18.7) | Ref | |
| Town/country | 111 (15.5) | 102 (20.3) | 0.74 (0.51-1.08) | 0.12 | 150 (12.8) | 0.87 (0.60-1.24) | 0.45 |
| Town | 413 (57.6) | 277 (55.2) | 0.97 (0.72-1.31) | 0.86 | **802 (68.5)** | **0.64 (0.49-0.84)** | $1.2\times10^{-3}$ |
| ***KRAS* rs712** | | | | | | | |
| G/G | 239 (33.3) | 160 (31.9) | Ref | | 364 (31.1) | Ref | |
| G/T | 341 (47.6) | 238 (47.4) | 0,87 (0,66-1,14) | 0.33 | 571 (48.8) | 0.83 (0.65-1.06) | 0.14 |
| T/T | 137 (19.1) | 104 (20.7) | 0,86 (0,61-1,21) | 0.38 | 236 (20.1) | 0.91 (0.67-1.24) | 0.55 |
| ***KIAA0182* rs709805** | | | | | | | |
| G/G | 367 (51.2) | 268 (53.4) | | | 622 (53.1) | | |
| G/A | 283 (39.5) | 206 (41.0) | | | 474 (10.5) | | |
| G/G + G/A | 650 (90.7) | 474 (94.4) | Ref | | 1096 (93.6) | Ref | |
| A/A | 67 (9.3) | **28 (5.6)** | **1.72 (1.06-2.78)** | $2.8\times10^{-2}$ | **75 (6.4)** | **1.57 (1.05-2.33)** | $2.7\times10^{-2}$ |
| ***IGSF22* rs2289965** | | | | | | | |
| A/A | 301 (42.0) | 211 (42.0) | Ref | | 458 (39.1) | Ref | |
| G/A | 323 (45.0) | 232 (46.2) | 0,97 (0,75-1,26) | 0.83 | 550 (47.0) | 0.94 (0.74-1.18) | 0.60 |
| G/G | 93 (13.0) | 59 (11.8) | 1,10 (0,75-1,64) | 0.62 | 163 (13.9) | 1.00 (0.71-1.41) | 0.99 |
| ***CD109* rs3012518** | | | | | | | |
| G/G | 284 (39.6) | 202 (40.2) | Ref | | 486 (41.5) | Ref | |
| G/A | 345 (48.1) | 236 (47.0) | 1,03 (0,79-1,33) | 0.83 | 520 (44.4) | 1.11 (0.88-1.40) | 0.38 |
| A/A | 88 (12.3) | 64 (12.8) | 0,99 (0,67-1,47) | 0.98 | 165 (14.1) | 0.96 (0.68-1.35) | 0.82 |
| ***PKNOX1* rs2839629** | | | | | | | |
| G/G | 224 (31.2) | 161 (32.1) | Ref | | 372 (31.8) | Ref | |
| G/A | 365 (50.9) | 242 (48.2) | 1,06 (0,80-1,39) | 0.70 | 568 (48.5) | 1.04 (0.82-1.33) | 0.75 |
| A/A | 128 (17.9) | 99 (19.7) | 0,96 (0,67-1,36) | 0.81 | 231 (19.7) | 0.93 (0.68-1.28) | 0.65 |
| ***ADARB2* rs904960** | | | | | | | |
| T/T | 336 (46.9) | 216 (43.0) | Ref | | 532 (45.5) | Ref | |
| T/C | 306 (42.7) | 234 (46.6) | 0,88 (0,68-1,13) | 0.32 | 524 (44.7) | 0.94 (0.75-1.18) | 0.59 |
| C/C | 75 (10.5) | 52 (10.4) | 0,94 (0,62-1,43) | 0.77 | 115 (9.8) | 0.96 (0.66-1.39) | 0.83 |
| ***EYA4* rs3734279** | | | | | | | |
| C/C | 261 (36.4) | 182 (36.3) | Ref | | 458 (39.1) | Ref | |
| C/T | 345 (48.1) | 245 (48.8) | 1,03 (0,79-1,35) | 0.82 | 545 (46.5) | 1.16 (0.92-1.47) | 0.21 |
| T/T | 111 (15.5) | 75 (14.9) | 1,58) | 0.65 | 168 (14.3) | 1.17 (0.84-1.64) | 0.36 |

(Continued)

**Table 4.** Characteristics of Colorectal Cancer Patients and Control Subjects and Adjusted Odds Ratio (OR) and 95% Confidence Interval (CI) Following Multivariate Logistic Regression Analysis (Continued)

| Characteristic | Cases n (%) | CFCCs n (%) | Adjusted OR (95% CI) | *P* | CFCC+HBDV n (%) | Adjusted OR (95% CI) | *P* |
|---|---|---|---|---|---|---|---|
| ***NUP210* rs354476** | | | | | | | |
| T/T | 195 (27.2) | 140 (27.9) | | | 318 (27.2) | | |
| T/C | 358 (49.9) | 272 (54.2) | | | 605 (51.6) | | |
| T/T+T/C | 553 (77.1) | 412 (82.1) | Ref | | 923 (78.8) | Ref | |
| C/C | 164 (22.9) | **90 (17.9)** | **1.36 (1.02-1.82)** | **$4.5\times10^{-3}$** | 248 (21.2) | 1.15 (0.88-1.49) | 0.30 |
| ***ABCB11* rs495714** | | | | | | | |
| A/A | 201 (28.0) | 162 (32.3) | Ref | | 329 (28.1) | Ref | |
| A/G | 357 (49.8) | 230 (45.8) | 1.24 (0.94-1.65) | 0.13 | 575 (49.1) | 1.09 (0.85-1.41) | 0.50 |
| G/G | 159 (22.2) | 110 (21.9) | 1.15 (0.82-1.62) | 0.42 | 267 (22.8) | 1.05 (0.77-1.43) | 0.76 |
| **Total** | **717** | **502** | | | **1171** | | |

Statistically significant results are shown in bold. Only the best genetic model is given (recessive for rs354476 *NUP210* and rs709805 *KIAA00182*).
BMI indicates body mass index; CFCC, cancer-free colonoscopy inspected controls; HBDV, healthy blood donor volunteers.

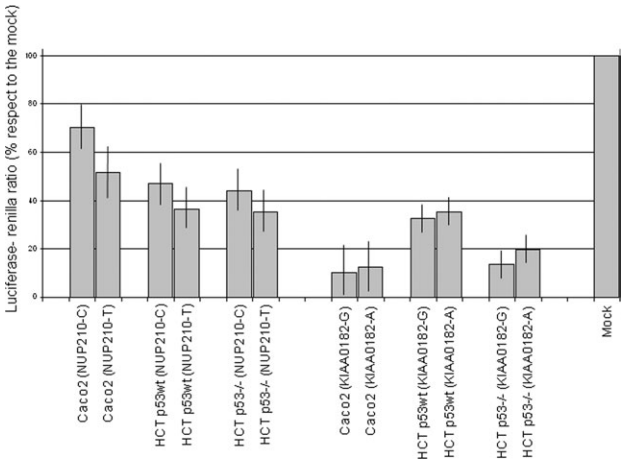


**Figure 1.** Luciferase/*Renilla* luminescence ratio is expressed as percent of the luminescence of the empty vector (EV), within each experiment. Six independent replicates in 3 independent experiments are summarized as mean values (± standard errors). The 3 cell lines employed in each experiment consistently show a reduced expression of luciferase, and the multifactor analysis of variance, considering all the results and cell lines together, shows a statistically significant difference between rs354476-C and rs354476-T ($P = .0037$).

CFCCs were pooled with the HBDV control group, the statistically significant association between the carriers of the rare allele for the SNP rs709805 within *KIAA0182* and the risk of CRC was confirmed. In this case, the adjusted OR for the rare homozygotes is 1.57 (95% CI = 1.05-2.33; $P = 2.7 \times 10^{-2}$). A negative familial history of CRC was also found to be associated with a reduced risk of CRC (OR = 0.56; 95% CI = 0.41-0.76), a trend (not statistically significant) observed also when only CFCCs were used.

Because 2 genotypes were associated with the risk of CRC, we carried out in vitro assays to investigate whether these SNPs could have some direct functional role in the regulation of expression (ie, translation) of *KIAA0182* and *NUP210*. For *KIAA0182*, we assayed rs709805, whereas for *NUP210*, we tested rs354476. The pmiR-GLO carrying both luciferase reporter gene and *Renilla* reference gene was chimerized by placing either the common or the variant form of the *KIAA0182* and *NUP210* 3′-UTRs at the 3′ end of the luciferase gene. Thus, the measurements of luminescence for luciferase were indicative of the intensity of its expression, depending on the 3′-UTR that was adopted. These measurements were compared to the internal reference (*Renilla*). Because the same vector carries both luciferase and *Renilla* genes, the experimental variability was greatly reduced compared with the typical experiments where 2 independent vectors are cotransfected. This allowed a more precise evaluation of the slight differences in the biological activity of the tested alleles. The background measurements of the luminescence of luciferase and *Renilla* were subtracted from the values obtained after transfection, and their ratio (luciferase/*Renilla*) was compared to the EV (vector not chimerized). Three cell lines derived from CRC were used: Caco2, HCT p53wt (wild-type for p53), and HCT p53−/− (lacking a functional p53). The results from 3 independent experiments (with 6 replicates for each point) are reported in Figure 1. Each luciferase/*Renilla* ratio is given as percent of the maximal intensity obtained with the EV (luciferase without 3′-UTR) within each experiment. In fact, at least in these series of experiments, the addition of a 3′-UTR at the luciferase gene led to a significant reduction of the luciferase expression. This could be due to a less stable messenger RNA, to the presence of negative regulators acting on the chimeric 3′-UTR

(eg, miRNAs), or to some other unknown mechanisms. The C-to-T point mutation of rs354476 (*NUP210*) resulted in a reduction of the expression of luciferase in all 3 cell lines employed. The average luciferase expression with the common C allele of *NUP210* was 53%, whereas it was 41.1% with the T allele, as compared to the EV. The MANOVA showed that the differences of the ratios between the C and the T alleles obtained by combining all the 3 cell lines are statistically significant ($P = .0035$). For *KIAA0182*, we observed a modest increase of the variant A allele (22.6%) compared with the common G allele (18.7% respect to the EV), but the ratios were not statistically significant (MANOVA, $P = .378$).

## DISCUSSION

In this work, we combined the information provided by the CRC mutome[20] with those from the microRNAome,[21] and we extracted 12 SNPs that have a potential role in affecting the regulation of 9 candidate genes (namely *ABCB11*, *ADARB2*, *KRAS*, *KIAA0182*, *IGSF22*, *CD109*, *NUP210*, *PKNOX1*, and *EYA4*), thereby potentially affecting the individual risk of CRC. Actually, we found an increased risk for the rs709805 within the *KIAA0182* gene in a population of cases and controls (either CFCCs or HBDVs+CFCCs) from the Czech Republic, the country with the highest incidence of CRC worldwide. A further association between rs354476 within *NUP210* and risk of CRC was significant only when considering the CFCC group. In this study, 2 different control populations were chosen. The inclusion of colonoscopy-negative individuals as controls (CFCCs) ensured cancer-free control individuals, because the negative result of colonoscopy serves as best available proof of CRC absence. Because the selection of these controls may not necessarily represent the healthy general population, we decided to expand the group by also including healthy individuals recruited from blood donor centers (HBDVs). However, it should be stressed that the HBDVs differ from cases for several covariates. There is a meaningful share of people with a younger age and a higher educational level than cases and HBDVs were mainly from the Prague district, whereas cases were collected from throughout the country. Thus, there could be explanations why the association with the SNP rs354476 within *NUP210* was not confirmed in the combined analysis (still maintaining the same trend). When cases were compared with the CFCCs, who share more similarities with CRC cases, the association between *NUP210* and CRC was strong enough to survive to the Bonferroni correction. Thus, it could be speculated that *NUP210* constitutes a risk factor for CRC when other preexisting predisposing conditions, such as inflammation or organ dysfunctions, are present.

In order to further interpret our results, we questioned whether the 9 candidate genes were actively transcribed in the normal colonic mucosa. In fact, the selection of these "CAN-genes" by Wood et al[20] was based on the resequencing of exonic regions, regardless the actual expression of the genes in the colon. By browsing the database of the University of Tokyo, Japan (http://www.lsbm.org/site_e/database/index.html), we found that only *KRAS*, *KIAA0182*, and *NUP210* are significantly expressed in the normal colonic mucosa. It is remarkable to note that 2 of 3 SNPs that compiled all the criteria related with the mutome, microRNAome, and transcriptome were associated with the risk of CRC. When the predicted biological effect (ie, the $|\Delta\Delta G|_{tot}$ of Table 2) was examined, the rs709805 within *KIAA0182* ranks as fourth, whereas the rs354476 (*NUP210*) ranks as eighth, but they ascended in rank to the second and third place, respectively, after filtering for the transcriptome. In vitro assays carried out to test the differences between the common and variant 3′-UTRs of *NUP210* and *KIAA0182* showed that the T allele of rs354476 (*NUP210*) was associated with a reduced expression of the reporter gene. This preliminary evidence does not prove any role of the predicted miRNAs in a differential regulation of *NUP210*. However, this finding highlights the fact that the alternative 3′-UTRs, placed into a "normal" cellular context and exploiting the "normal" cellular machineries, have different capacities in determining the levels of expression of *NUP210*. We employed colorectal cell lines, and it is conceivable that these cell lines express a set of miRNAs similar to that of normal colorectal cells. However, we cannot say whether miRNAs, messenger RNA stability, or other mechanisms are involved in the observed genotype-dependent differential expression of *NUP210*. No clear evidence came from the assay carried out on the *KIAA0182* 3′-UTR. *KIAA0182* maps at 16q24.1 and, as suggested by its designation, it belongs to a family of more than 2000 genes. It encodes a putative genetic suppressor element 1 protein and it might exhibit RNA-binding activity.[33] However, there is little information about this gene, whereas more is known about *NUP210*. This gene maps at 3p25.1 and encodes the nuclear pore glycoprotein-210 (gp210) involved in the structural organization of the nuclear pore complex.[34] Interestingly, during mitosis, Ser1880 of gp210 is phosphorylated by the cyclin B-p34cdc2[35] and an increased expression of *NUP210* was found in cervical cancer.[36]

According to The Roche Cancer Genome Database (http://rcgdb.bioinf.uni-sb.de/MutomeWeb/), somatic mutations within *KIAA0182* and *NUP210* were already described not only for CRC (*KIAA0182* = c.366_dupC and c.1879_C>T encoding for p.R627W; *NUP210* = c.2951G>A encoding for R984H; c.923 C>T encoding for S308L; IVS6-3C>T) but also for malignant melanoma (*KIAA0182* = c.2172C>T; *NUP210* = c.3447C>T). The presence of somatic mutations in the same genes in different types of cancer reinforces the hypothesis that these genes play an important role in human carcinogenesis. *NUP210* and *KIA0182* were also proposed among the biomarkers for human CRC (eg, see www.wipo.int/patentscope/search/en/WO2006081248). According to our results, it is likely that functional SNPs could modulate the normal levels of these proteins. Therefore, knowledge of the effects of the SNPs is also very important to appropriately set the correct thresholds to distinguish normal concentrations from pathological ones.

In conclusion, this hypothesis-driven study carried out using all the latest "-omics" information available from the literature suggests for the first time that the regulation of *NUP210* and *KIAA0182* may be important for modulating the risk of CRC. Future work is warranted to validate the results in other populations and to explore further the biological significance of the mentioned SNPs.

## FUNDING SOURCES

This work was supported by AIRC (Associazione Italiana Ricerca Cancro, investigator grant year 2008), by the MIUR (Italian Ministry of Research, PRIN) and the University of Pisa (Ex60%) and was supported by the Grant Agency of the Czech Republic: CZ:GACR:GA P304/10/1286, CZ:GACR:GA 310/07/1430, and CZ:GACR:GA 305/09/P194.

## CONFLICT OF INTEREST DISCLOSURE

The authors made no disclosure.

## REFERENCES

1. John BJ, Irukulla S, Abulafi AM, Kumar D, Mendall MA. Systematic review: adipose tissue, obesity and gastrointestinal diseases. *Aliment Pharmacol Ther*. 2006;23:1511-1523.
2. Ferguson LR. Meat and cancer. *Meat Sci*. 2010;84:308-313.
3. Sinicrope FA, Foster NR, Sargent DJ, O'Connell MJ, Rankin C. Obesity is an independent prognostic variable in colon cancer survivors. *Clin Cancer Res*. 2010;16:1884-1893.
4. Tuynman JB, Peppelenbosch MP, Richel DJ. COX-2 inhibition as a tool to treat and prevent colorectal cancer. *Crit Rev Oncol Hematol*. 2004;52:81-101.
5. Kim MK, Park JH. Conference on "Multidisciplinary approaches to nutritional problems". Symposium on "Nutrition and health". Cruciferous vegetable intake and the risk of human cancer: epidemiological evidence. *Proc Nutr Soc*. 2009;68:103-110.
6. Broderick P, Carvajal-Carmona L, Pittman AM, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat Genet*. 2007;39:1315-1317.
7. Tenesa A, Farrington SM, Prendergast JG, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat Genet*. 2008;40:631-637.
8. Tomlinson I, Webb E, Carvajal-Carmona L, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet*. 2007;39:984-988.
9. Zanke BW, Greenwood CM, Rangrej J, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet*. 2007;39:989-994.
10. Houlston RS, Cheadle J, Dobbins SE, et al. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat Genet*. 2010;42:973-977.
11. Nannya Y, Taura K, Kurokawa M, Chiba S, Ogawa S. Evaluation of genome-wide power of genetic association studies based on empirical data from the HapMap project. *Hum Mol Genet*. 2007;16:2494-2505.
12. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev*. 2009;19:212-219.
13. Gao Y, Cao Y, Tan A, Liao C, Mo Z, Gao F. Glutathione S-transferase M1 polymorphism and sporadic colorectal cancer risk: An updating meta-analysis and HuGE review of 36 case-control studies. *Ann Epidemiol*. 2010;20:108-121.
14. Tian Y, Li Y, Hu Z, Wang D, Sun X, Ren C. Differential effects of NOD2 polymorphisms on colorectal cancer risk: a meta-analysis. *Int J Colorectal Dis*. 2010;25:161-168.
15. Hu Z, Li X, Qu X, et al. Intron 3 16 bp duplication polymorphism of TP53 contributes to cancer susceptibility: a meta-analysis. *Carcinogenesis*. 2010;31:643-647.
16. Cai Y, Yu X, Hu S, Yu J. A brief review on the mechanisms of miRNA regulation. *Genomics Proteomics Bioinformatics*. 2009;7:147-154.
17. Tan Z, Randall G, Fan J, et al. Allele-specific targeting of microRNAs to HLA-G and risk of asthma. *Am J Hum Genet*. 2007;81:829-834.
18. Bao BY, Pao JB, Huang CN, et al. Polymorphisms inside microRNAs and microRNA target sites predict clinical outcomes in prostate cancer patients receiving androgen-deprivation therapy. *Clin Cancer Res*. 2011;17:928-936.
19. Landi D, Moreno V, Guino E, et al. 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*. 2011;717:109-115.
20. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. *Science*. 2007;318:1108-1113.
21. Cummins JM, He Y, Leary RJ, et al. The colorectal microRNAome. *Proc Natl Acad Sci U S A*. 2006;103:3687-3692.
22. Pardini B, Naccarati A, Novotny J, et al. DNA repair genetic polymorphisms and risk of colorectal cancer in the Czech Republic. *Mutat Res*. 2008;638:146-153.
23. Pardini B, Naccarati A, Polakova V, et al. NBN 657del5 heterozygous mutations and colorectal cancer risk in the Czech Republic. *Mutat Res*. 2009;666:64-67.
24. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*. 2006;34:D140-D144.
25. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human microRNA targets. *PLoS Biol*. 2004;2:e363.
26. Rusinov V, Baev V, Minkov IN, Tabler M. MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence. *Nucleic Acids Res*. 2005;33:W696-W700.
27. Kiriakidou M, Nelson PT, Kouranov A, et al. A combined computational-experimental approach predicts human microRNA targets. *Genes Dev*. 2004;18:1165-1178.

28. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell*. 2003;115: 787-798.
29. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. *Nucleic Acids Res*. 2008;36: W70-W74.
30. Landi D, Gemignani F, Naccarati A, et al. Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. *Carcinogenesis*. 2008;29:579-584.
31. Landi D, Barale R, Gemignani F, Landi S. Prediction of the biological effect of polymorphisms within microRNA binding sites. *Methods Mol Biol*. 2011;676:197-210.
32. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559-575.
33. Nagase T, Seki N, Ishikawa K, Tanaka A, Nomura N. Prediction of the coding sequences of unidentified human genes. V. The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res*. 1996;3:17-24.
34. Greber UF, Senior A, Gerace L. A major glycoprotein of the nuclear pore complex is a membrane-spanning polypeptide with a large lumenal domain and a small cytoplasmic tail. *EMBO J*. 1990;9:1495-1502.
35. Favreau C, Worman HJ, Wozniak RW, Frappier T, Courvalin JC. Cell cycle-dependent phosphorylation of nucleoporins and nuclear pore membrane protein Gp210. *Biochemistry*. 1996;35: 8035-8044.
36. Rajkumar T, Sabitha K, Vijayalakshmi N, et al. Identification and validation of genes involved in cervical tumourigenesis. *BMC Cancer*. 2011;11:80.