Tohoku J. Exp. Med., 2020, **250**, 109-119

CD44 Genotypes Are Associated with Susceptibility and Tumor Characteristics in Colorectal Cancer Patients

Rosa María Márquez-González,[1] Anilú Margarita Saucedo-Sariñana,[1] Patricio Barros-Núñez,[2] Martha Patricia Gallegos-Arreola,[3] Tomás Daniel Pineda-Razo,[4] María Eugenia Marin-Contreras,[5] Silvia Esperanza Flores-Martínez,[1] José Sánchez-Corona[1] and Mónica Alejandra Rosales-Reynoso[1]

[1]Molecular Medicine Division, Centro de Investigación Biomédica de Occidente (CIBO), Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco, Mexico
[2]Research Unit of Metabolic Diseases, Pediatric UMAE, Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco, Mexico
[3]Genetic Division, Centro de Investigación Biomédica de Occidente (CIBO), Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco, Mexico
[4]Medical Oncology Service, Specialty Hospital, Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco, Mexico
[5]Gastroenterology Service, Specialty Hospital, Instituto Mexicano del Seguro Social (IMSS), Guadalajara, Jalisco, Mexico

Colorectal cancer is the third cause of cancer and the second leading cause of death worldwide. The *CD44* gene plays a key role in malignant processes, including growth, survival, epithelial to mesenchymal transition and metastasis. It is also known that some variants as rs187116 (c.67+4883G>A) and rs7116432 (c.2024+779A>G) can modulate the function of the *CD44* gene and malignant transformation in several neoplasms. This study aims to explore, for the first time, the association of the *CD44* rs187116 and rs7116432 variants in patients with colorectal cancer. Genomic DNA from 250 patients and 250 healthy blood donors were analyzed. The identification of variants was made by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) methodology. Association was calculated by the odds ratio (OR) test and multivariate analysis. Individuals carrying the G/A and A/A genotypes for the rs187116 polymorphism showed an increased risk for colorectal cancer (OR = 3.11, 95% CI: 1.87-5.16, P = 0.001 and OR = 3.59, 95% CI: 2.06-6.25, P = 0.001, respectively). After adjusting for age and gender, these same genotypes and the G/G genotype of the rs7116432 polymorphism were associated with TNM stage and tumor location in the colon. Moreover, the A-G (rs187116 and rs7116432) haplotype was associated with increased risk; while, the haplotype G-A (rs187116 and rs7116432) was related with decreased risk. In conclusion, our results suggest that the here analyzed *CD44* variants are involved with risk, TNM stage and tumor location in colorectal cancer.

Keywords: colorectal cancer; *CD44*; haplotypes; polymorphisms; susceptibility
Tohoku J. Exp. Med., 2020 February, **250** (2), 109-119.

Introduction

Colorectal cancer (CRC) is the third most frequent cancer in American countries (Bray et al. 2018). In Mexico, the CRC incidence was 11,484 cases, with a rate of 9.6/100,000 inhabitants and a mortality rate of 4.1/100,000 inhabitants (World Health Organization 2018). CRC is a complex disease which is caused by environmental factors, genetic and epigenetic alterations in different kinds of genes (oncogenes, mismatch repair, cell cycle and tumor suppressor genes), which alter signaling pathways that have been identified as molecular mechanisms underlying to CRC

Received December 23, 2019; revised and accepted February 6, 2020. Published online February 28, 2020; doi: 10.1620/tjem.250.109.
Correspondence: Mónica Alejandra Rosales-Reynoso, Ph.D., División de Medicina Molecular, Centro de Investigación Biomédica de Occidente (CIBO), Instituto Mexicano del Seguro Social (IMSS), Sierra Mojada # 800, Colonia Independencia, Guadalajara, Jalisco CP 44340, México.
e-mail: mareynoso77@yahoo.com.mx or monica.rosales@imss.gob.mx
©2020 Tohoku University Medical Press. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC-BY-NC-ND 4.0). Anyone may download, reuse, copy, reprint, or distribute the article without modifications or adaptations for non-profit purposes if they cite the original authors and source properly.
https://creativecommons.org/licenses/by-nc-nd/4.0/

development (Fischer et al. 2019). When some of these genetic alterations and environmental factors accumulate, they can lead to the onset and progression of CRC through a multistage tumorigenesis process known as the adenoma-carcinoma sequence (Fearon and Vogelstein 1990). It is assumed that the tumorigenesis process is caused by dysregulation of crucial signaling pathways in cancer stem cells (CSCs), altering the balance of several cellular biological processes that include growth, proliferation and apoptosis (Orian-Rousseau and Schmitt 2015; Skandalis et al. 2019).

CSCs are a group of tumor cells with self-renewal potential, recognized as responsible for tumor development, chemoresistance, metastasis and recurrence (Hen and Barkan 2019; Rossari et al. 2019; Skandalis et al. 2019). These cells express various specific or non-specific CSCs markers; some of them have been utilized for cell separation and characterization (Atashzar et al. 2020). These markers increase in the advanced stages of many types of cancer and correlate with clinicopathological parameters, becoming valuable biomarkers for cancer diagnosis and prognosis (Atashzar et al. 2020).

Currently, some tumor-initiating cells markers as CD24, CD26, CD29, CD44; CD133, CD166; Musashi-1, Leucine-rich repeat-containing G-protein coupled receptor and the aldehyde dehydrogenase 1 family member A1 have been reported in CRC (Reya et al. 2001; Ricci-Vitiani et al. 2007; Dalerba et al. 2007; Vermeulen et al. 2008; Mueller et al. 2009; Sanders and Majumdar 2011). Among them, CD166 (ALCAM) correlates with overall survival, and it is considered a predictive biomarker for stage II in CRC patients (Han et al. 2017; Wahab et al. 2017). Likewise, the CD44 expression correlates 100% in CRC patients, and its expression is significantly associated with poor prognosis (Huh et al. 2009; Wahab et al. 2017). CD44 is a cell surface glycoprotein encoded by the *CD44* gene and plays a role in numerous cellular processes as cell-cell interactions, cell adhesion, migration and regulation of growth and homing of lymphocytes (Spring et al. 1988; Dalerba et al. 2007).

The effect of *CD44* polymorphisms on human cancer susceptibility, prognosis and recurrence has been assessed in several human cancers as gastric adenocarcinoma, breast, hepatocellular, prostate and colorectal cancer (Gerger et al. 2011; Winder et al. 2011; Zhou et al. 2011; Chou et al. 2014; Stremitzer et al. 2015; Stotz et al. 2017; Freedman et al. 2018; Wan et al. 2019). Among them, the rs187116 (c.67+4883G>A) polymorphism, located in intron 1 of the *CD44* gene, is the most frequently studied (Winder et al. 2011; Suenaga et al. 2015; Sapcharoen et al. 2019) and proposed as genetic marker that predicting the clinical outcome in cancer patients (Tulsyan et al. 2013; Liu et al. 2015; Lin et al. 2017; Tongtawee et al. 2017; Bitaraf et al. 2018). Another polymorphism in the *CD44* gene is rs7116432 (c.2024+779A>G) located in non-coding exon 19, which has been associated only with overall survival and tumor recurrence in gastric cancer (Bitaraf et al. 2018). Accurately, in CRC, the *CD44* rs187116 and rs7116432 polymorphisms have been analyzed only for prognosis and disease recurrence (Gerger et al. 2011; Stremitzer et al. 2015; Stotz et al. 2017). Although the biological significance of these polymorphisms remains unknown, it is plausible to think that the *CD44* polymorphisms may have a direct impact on the regulation of gene-splicing mechanism, explaining its participation in numerous tumorigenic processes. This study aims to examine for the first time the distribution of alleles, genotypes and haplotypes of these two polymorphisms of the *CD44* gene (rs187116 G>A and rs7116432 A>G), evaluating their possible association with the development and with clinicopathological characteristics of CRC.

Materials and Methods

Subjects

This study included 250 clinically diagnosed and histologically confirmed patients with sporadic colorectal adenocarcinoma between 2014 and 2018, according to the Clinical Practice Guidelines in the colon and rectal cancer and the clinicopathological criteria of the Specialty Hospital of the West National Medical Center of the Mexican Institute of Social Security (IMSS) in Guadalajara, Jalisco. Pathological tumor staging and grading were performed according to the tumor-node-metastasis (TNM) classification. The control group included 250 healthy unrelated individuals, colonoscopy-negative for malignancy and not matched by age with the patient group. The exclusion criteria used for the patients and controls were: autoimmune or inflammatory bowel disease, personal history of cancer, or familial history of any known hereditary cancer syndromes. The study was approved by the Ethical Committee 1305 (R-2014-1305-8 and R-2018-1305-001) of West Biomedical Research Center, IMSS, and conducted according to national and international ethical standards. All the participants signed the informed consent for participation in this study. An epidemiological questionnaire allowed us to collect personal data, including age, gender, drinking and smoking status and familial history. Information about the clinical and pathological features of the patients was obtained from the hospital records.

Genotyping

Genomic DNA was isolated from peripheral blood using standard methods (Miller et al. 1988). The polymorphisms rs187116 (G>A) and rs7116432 (A>G) in the *CD44* gene were genotyped by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) using the following primer pairs: rs7116432-F: 5’- CAT CGT CTT CTT GCT GTT AGG A-3’ and 7116432-R: 5’-GGT CTT GGT TCA GGT AGG GAG A-3’ and rs187116-F: 5’-AGG TGG TTG GAG ATC ACC TG -3’, rs187116-R: 5’-CTT TCG CAA GAA CCA CTT CC -3’ (Bitaraf et al. 2018). For the rs7116432 polymorphism, PCR was performed

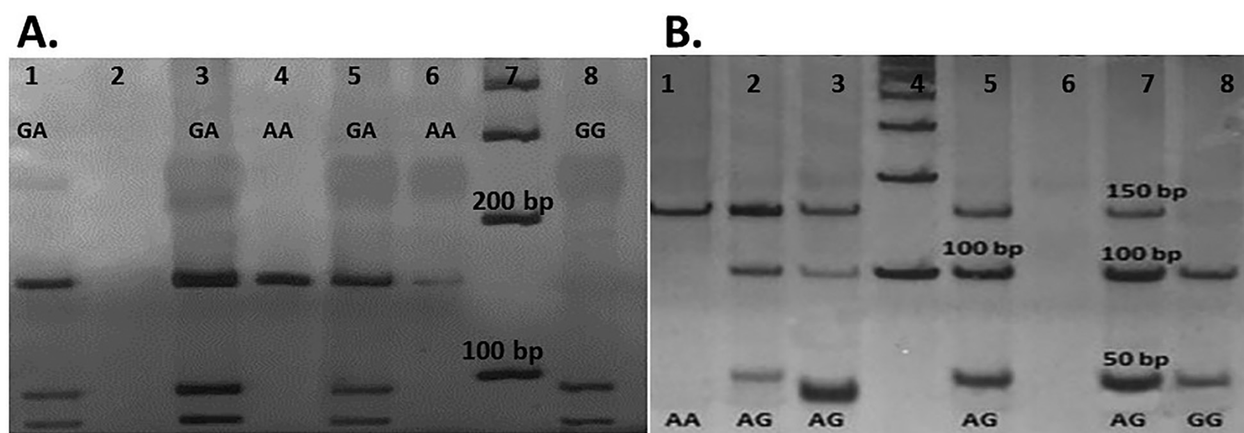

Fig. 1. Restriction fragment length analysis on 8% polyacrylamide gels stained with AgNO3.
A. Enzymatic digestion with the MspI endonuclease to identify the genotypes of the *CD44* rs187116 polymorphism. Lanes 1, 3, 5: Individuals with genotype heterozygous (G/A) showing bands of 153 + 93 + 60 bp. Lanes 4 and 6: Individuals with genotype homozygous polymorphic (A/A) showing band of 153 bp. Lane 8: Individual with genotype homozygous wildtype (G/G) showing bands of 93 + 60 bp. Lane 2: without sample.
B. Enzymatic digestion with NIaIII endonuclease to identify the genotypes of the *CD44* rs7116432 polymorphism. Lane 1: Individual with genotype homozygous wildtype (A/A) showing a band of 150 bp. Lanes 2, 3, 5, and 7: Individuals with genotype heterozygous (A/G) showing bands of 150 + 100 + 50 bp. Lane 6: without sample. Lane 8: Individual with genotype homozygous polymorphic (G/G) showing bands of 100 + 50 bp.

using a 10 μL volume containing, 10X buffer (500 mM KCl, 100 mM Tris- HCl and 0.1% Triton TMX-100), 2.0 mM MgCl2, 150 μM dNTPs, 1 μM of each primer, 2 U Taq DNA Polymerase and 100 ng DNA. Denaturation was carried out at 95°C for 5 min, annealing at 61°C for 30 sec, and elongation at 72°C for 30 sec and a final extension for 5 min at 72°C for 35 cycles. The rs187116 polymorphism was identified under the same PCR conditions except for the annealing temperature (58°C). Four units of MspI enzyme restriction at 37°C overnight (New England Biolabs, USA) were used to digest 5 μL of the PCR amplification product of the rs187116 polymorphism, according to the manufacturer's instruction. The digested products were separated on 8% polyacrylamide gels. The MspI endonuclease recognizes the restriction site CCGG and cuts between the 2 cytosines. The homozygous wildtype genotype G/G contains a recognition site for the enzyme MspI, so yields two DNA fragments of 93 and 60 bp in length. The homozygous polymorphic genotype A/A does not contain a recognition site for the enzyme MspI, so the 153 bp amplificon remains unaltered after incubation with MspI enzyme. Digestion of the amplificon containing the heterozygous genotype G/A produces three DNA fragments of 153, 93, and 60 bp (Fig. 1A). For the rs7116432 polymorphism, five microliters of the PCR product were digested with 3U of NlaIII restriction enzyme at 37°C overnight (New England Biolabs, USA), according to the manufacturer's instructions, and separated on 8% polyacrylamide gels. The NlaIII endonuclease recognizes the restriction site CATG and cuts after the guanine nucleotide. The homozygous wildtype genotype A/A does not contain a recognition site for the enzyme NlaIII, so the 150 bp amplificon remains unaltered after incubation with NlaIII. The homozygous polymorphic genotype G/G contains a recognition site for the enzyme NlaIII, so the digestion yields two DNA fragments of 100 and 50 bp. Lastly, the heterozygous genotype A/G produces three DNA fragments 150, 100 and 50 bp (Fig. 1B). The quality control for these assays was evaluated by re-genotyping 10% of samples, randomly selected by an independent technician. Concordance between genotyping tests was 100%.

Statistical analysis

Genotype and allele frequencies were estimated by direct counting in both groups. The Chi-square test assessed the Hardy-Weinberg equilibrium (HWE). The chi-square test also evaluated differences in genotype and allele distributions with the clinical features of patients and controls. Statistical analysis included odds ratio analysis and Yates corrected Chi-square test. Association of genotypes or alleles with CRC and stratified TNM stage were calculated by odds ratio (OR) and confidence intervals (CI) in an SPSS v17.0 software package (SPSS Inc., Chicago, IL, USA). Logistic regression was then used for multivariate analyses employing SPSS software in order to include demographic, clinical data, and confounding variables such as gender and age. Haplotype analysis was performed using the Haploview 4.2 software. For all statistical analyses, $P < 0.05$ was considered significant.

Results

Characteristics of the subjects included in the study

The CRC group included 250 patients (145 males and 105 females); meanwhile, the control group included 117 males and 133 females (250 healthy individuals), all of them efficiently genotyped for the *CD44* rs187116 and

Table 1. Demographic and clinical characteristics of the CRC patients and control subjects.

Characteristic	CRC group n = 250 (%)	Control group n = 250 (%)	P value
Mean Age (years)			
Mean	58.6	38.2	**0.001**
< 50	54 (21.6)	202 (80.8)	
> 50	196 (78.4)	48 (19.2)	**0.001**
Sex			
Female	105 (42)	133 (53.2)	0.151
Male	145 (58)	117 (46.8)	
Smoking status			
Yes	87 (35)	89 (35.6)	0.851
No	163 (65)	161 (64.4)	
Drinking status			
Yes	73 (29)	70 (28)	0.766
No	177 (71)	180 (72)	
Family History			
- Diabetes Mellitus	26 (10)	86 (34)	
- Hypertension	18 (7.2)	84 (35)	0.078
TNM pathological stage			
I	6 (2)	-	
II	70 (28)	-	
III	94 (38)	-	
IV	80 (32)	-	
Tumor location			
Colon	140 (56)	-	
Rectum	110 (44)	-	

P values were calculated by the Chi-square test.

rs7116432 polymorphisms. Table 1 shows a comparative analysis of demographic and clinical data from CRC patients and control subjects. Significant differences in age distribution were detected (P = 0.001). The CRC group showed an average age of 58.6 years (range 18 to 91); for the control group, the mean age was 38.2 years (range 29 to 59). There were no significant differences between the two groups respecting gender, family history and smoking and drinking consumption.

Genotype frequencies of the CD44 polymorphisms in CRC patients and control individuals

The comparative analysis of both *CD44* variants in CRC patients and controls showed significant differences (Table 2). In the control group, the two analyzed SNPs were observed in Hardy-Weinberg equilibrium (data not shown). The G/G wildtype genotype of the rs187116 polymorphism was found in 10.8% (27/250) of CRC patients and 28.4% (71/250) of control group individuals. The G/A genotype was observed in 56.4% (141/250) of CRC patients and 48% (119/250) of control group, showing that the G/A genotype is positively associated to CRC (OR = 3.11; 95% CI = 1.87-5.16, P = 0.001). The A/A polymorphic genotype was observed in 32.8% (82/250) of CRC patients and 24% (60/250) of control individuals, revealing that the A/A genotype was also associated with CRC (OR = 3.59; 95% CI = 2.06-6.25, P = 0.001). Under a dominant pattern of allelic interaction (G/A+A/A vs. G/G), the allele A was also correlated to CRC (OR = 3.27; 95% CI = 2.01-5.32, P = 0.001). Allelic frequencies were also significantly different, demonstrating that carriers of the A allele have increased the risk of CRC (OR = 1.70; 95% CI = 1.32-2.19, P = 0.001).

Regarding the polymorphism *CD44* rs7116432, the A/A wildtype genotype was observed in 28% (71/250) of CRC patients and 40.8% (102/250) of control individuals. The A/G genotype was observed in 46% (115/250) of CRC patients and 47% (118/250) of control individuals. The G/G polymorphic genotype was observed in 26% (64/250) of CRC patients and 12% (30/250) of control individuals, exhibiting a significant difference (OR = 3.06; 95% CI = 1.80-5.20, P = 0.001). Under a dominant pattern of allelic interaction, the risk analysis A/G+G/G vs. A/A showed an association between the allele G and CRC risk (OR = 1.80; 95% CI = 1.24-2.61, P = 0.002). Likewise, the allele frequencies were different (OR = 1.71; 95% CI = 1.32-2.20, P = 0.001); showing the allele G is associated with CRC.

Table 2. Distribution of genotypes and allelic frequencies of the *CD44* gene polymorphisms and the risk of CRC.

Genotype	Frequencies: CRC group n = 250 (%)	Frequencies: Control group n = 250 (%)	OR (95% CI)	P value
***CD44* (rs187116)**				
G/G	27(10.8)	71 (28.4)	1.00 (Reference)	
G/A	141 (56.4)	119 (48)	**3.11 (1.87-5.16)**	**0.001**
A/A	82 (32.8)	60 (24)	**3.59 (2.06-6.25)**	**0.001**
G/A+A/A vs. G/G	223 (89.2)	179 (71.6)	**3.27 (2.01-5.32)**	**0.001**
Allele				
G	195 (0.39)	261 (0.48)	1.00 (Reference)	
A	305 (0.61)	239 (0.52)	**1.70 (1.32-2.19)**	**0.001**
***CD44* (rs7116432)**				
A/A	71 (28)	102 (40.8)	1.00 (Reference)	
A/G	115 (46)	118 (47)	1.40 (0.94-2.82)	0.118
G/G	64 (26)	30 (12)	**3.06 (1.80-5.20)**	**0.001**
A/G+G/G vs. A/A	179 (72)	148 (59.2)	**1.80 (1.24-2.61)**	**0.002**
Allele				
A	257 (0.51)	322 (0.64)	1.00 (Reference)	
G	243 (0.49)	178 (0.35)	**1.71 (1.32-2.20)**	**0.001**

P values were calculated by chi-square test.

Table 3. Stratification analysis of the association of *CD44* polymorphisms with CRC under genetic models.

rs187116

Variable	Case/Control GG	Case/Control GA	Case/Control AA	OR (95% CI); P value GA vs. GG	OR (95% CI); P value AA vs. GG	OR (95% CI); P value GA+AA vs. GG
Sex						
Male	13/35	86/56	46/26	**4.13 (2.01-8.49); 0.001**	**4.76 (2.14-10.5); 0.001**	**4.33 (2.16-8.67); 0.001**
Female	14/36	55/63	36/34	**2.44 (1.09-4.59); 0.038**	**2.72 (1.25-5.91); 0.017**	**2.41 (1.22-4.76); 0.015**
Smoking						
Yes	4/12	51/57	32/20	2.68 (0.81-8.84); 0.161	**4.80 (1.35-16.9); 0.022**	3.23 (1.00-10.4); 0.073
No	23/28	90/92	50/41	1.19 (0.63-2.22); 0.695	1.48 (0.74-2.95); 0.341	1.28 (0.70-2.33); 0.510
Alcohol						
Yes	5/9	46/41	22/20	2.01 (0.62-6.51); 0.366	1.98 (0.56-6.90); 0.440	2.00 (0.63-6.31); 0.353
No	22/42	95/100	60/38	0.92 (0.48-1.73); 0.930	1.79 (0.88-3.62); 0.144	1.13 (0.61-2.10); 0.801
Age (years)						
< 50	16/59	27/97	11/46	1.02 (0.51-2.06); 1.000	2.35 (0.80-6.83); 0.180	0.88 (0.37-2.08); 0.944
> 50	21/12	104/22	71/14	**2.70 (1.16-6.29); 0.034**	**2.89 (1.16-7.21); 0.036**	**2.77 (1.25-6.15); 0.018**

rs7116432

Variable	Case/Control AA	Case/Control AG	Case/Control GG	OR (95% CI); P value AG vs. AA	OR (95% CI); P value GG vs. AA	OR (95% CI); P value AG+GG vs. AA
Sex						
Male	35/44	69/63	41/10	1.37 (0.78-2.41); 0.327	**5.15 (2.26-11.7); 0.001**	**1.89 (1.11-3.22); 0.025**
Female	36/58	46/55	23/20	1.34 (0.76-2.38); 0.379	1.85 (0.89-3.84); 0.138	1.48 (0.87-2.51); 0.184
Smoking						
Yes	23/20	42/50	22/19	0.73 (0.35-1.51); 0.506	1.00 (0.42-2.37); 1.000	0.80 (0.40-3.76); 0.662
No	48/63	73/71	42/27	1.34 (0.82-2.21); 0.291	2.04 (1.10-3.76); 0.301	1.54 (0.97-2.44); 0.085
Alcohol						
Yes	20/19	36/35	17/16	0.97 (0.44-2.13); 1.000	1.00 (0.39-2.55); 1.000	0.98 (0.47-2.06); 1.000
No	51/73	79/81	47/26	1.39 (0.86-2.24); 0.206	2.58 (1.42-4.70); 0.202	1.68 (1.00-2.61); 0.065
Age (years)						
< 50	12/75	22/107	20/20	1.28 (0.59-2.75); 0.649	**6.25 (2.62-14.90); 0.001**	1.91 (0.95-3.86); 0.093
> 50	59/17	88/25	49/6	1.01 (0.50-2.04); 1.000	2.35 (0.86-6.42); 1.141	1.27 (0.65-2.47); 0.590

Association of the CD44 genotypes with gender, age, alcohol and smoking consumption

The analysis of each SNP by gender, age, consumption of alcohol and smoking are shown in Table 3. CRC males shown a significant association with the G/A and A/A genotypes of the rs187116 polymorphism (OR = 4.13; 95% CI = 2.01-8.49, P = 0.001 and OR = 4.76; 95% CI = 2.14-10.57, P = 0.001, respectively). Such an association was also evident under the dominant model of inheritance G/A+A/A vs. G/G (OR = 4.33; 95% CI = 2.16-8.67, P = 0.001). Regarding to age, the CRC patients > 50 years shown an increased risk in presence of G/A, and A/A genotypes (OR = 2.70; 95% CI = 1.16-6.29, P = 0.034 and OR = 2.89; 95% CI = 1.16-7.21, P = 0.036, respectively). This association was also observed under the dominant model of inheritance G/A+A/A vs. G/G (OR = 2.77; 95% CI = 1.25-6.15, P = 0.018). Regarding the location of the tumor, an association was found only for tobacco in patients carrying the A/A genotype (OR = 4.80; 95% CI = 1.35-16.9, P = 0.022).

For the rs7116432 polymorphism, CRC males were associated with the G/G genotype (OR = 5.15; 95% CI = 2.26-11.72, P = 0.001) and this association was also evident under the dominant model of inheritance (OR = 1.89; 95% CI = 1.11-3.22, P = 0.025). Regarding the age, we observed an increased risk in patients < 50 years carrying the G/G genotype (OR = 6.25; 95% CI = 2.62-14.90, P = 0.001). With respect to the tobacco and alcohol consumption, we did not observe any significant difference.

Association of CD44 genotypes with TNM stage and tumor location

Analysis of each SNP by TNM stage and tumor loca-

Table 4. Association between TNM stage and the genotypes distribution for the rs187116 and rs7116432 polymorphisms of the *CD44* gene.

SNP rs187116 Genotype	I+II stage n = 76 (%)	III+IV Stage n = 174 (%)	Control n = 250 (%)	I+II Stage vs. Control OR (95% CI)	P value	I+II Stage vs. Control OR (95% CI)*	P value*	III+IV Stage vs. Control OR (95% CI)	P value	III+IV Stage vs. Control OR (95% CI)*	P value*
G/G	6 (7)	21 (12)	71 (28)	1.00 (Reference)		1.00 (Reference)		1.00 (Reference)		1.00 (Reference)	
G/A	46 (61)	95 (55)	119 (48)	**4.57 (1.85-11.2)**	**0.001**	**4.12 (1.54-11.0)**	**0.005**	**2.69 (1.54-4.70)**	**0.001**	**2.44(1.26-4.74)**	**0.008**
A/A	24 (32)	58 (33)	60 (24)	**4.73 (1.81-12-3)**	**0.001**	**3.59 (1.25-10.3)**	**0.017**	**3.26 (1.78-5.99)**	**0.001**	1.97 (0.90-4.30)	0.088
G/A+A/A	70 (92)	153 (88)	179 (72)	**4.62 (1.92-11.1)**	**0.001**	**3.94 (1.52-10.1)**	**0.005**	**2.88 (1.67-4.92)**	**0.001**	**2.31 (1.21-4.40)**	**0.010**

SNP rs7116432 Genotype	I+II stage n = 76 (%)	III+IV Stage n = 174 (%)	Control n = 250 (%)	I+II Stage vs. Control OR (95% CI)	P value	I+II Stage vs. Control OR (95% CI)*	P value*	III+IV Stage vs. Control OR (95% CI)	P value	III+IV Stage vs. Control OR (95% CI)*	P value*
A/A	21 (28)	50 (29)	102 (41)	1.00 (Reference)		1.00 (Reference)		1.00 (Reference)		1.00 (Reference)	
A/G	36 (47)	79 (45)	118 (47)	1.48 (0.81-2.69)	0.254	1.27 (0.64-2.54)	0.487	1.36 (0.87-2.12)	0.203	1.50 (0.86-2.62)	0.145
G/G	19 (25)	45 (26)	30 (12)	**3.07 (1.46-6.46)**	**0.004**	**2.91 (1.22-6.92)**	**0.015**	**3.06 (1.72-5.42)**	**0.001**	**3.36 (1.51-7.49)**	**0.003**
AG/+G/G	55 (72)	124 (71)	148 (59)	**1.80 (1.02-3.16)**	**0.000**	1.56 (0.82-2.97)	0.169	**1.70 (1.12-2.58)**	**0.014**	**1.82 (1.08-3.05)**	**0.024**

Bold values are statistically significant (P < 0.05).
*Adjust for age and sex.

Table 5. Association between tumor location and the genotypes distribution for the rs187116 and rs7116432 polymorphisms of the *CD44* gene.

SNP rs187116 Genotype	Colon n = 140 (%)	Rectum n = 110 (%)	Control n = 250 (%)	Colon vs. Control OR (95% CI)	P value	Colon vs Control OR (95% CI)*	P value*	Rectum vs. Control OR (95% CI)	P value	Rectum vs Control OR (95% CI)*	P value*
G/G	11 (8)	16 (14)	71 (28)	1.00 (Reference)		1.00 (Reference)		1.00 (Reference)		1.00 (Reference)	
G/A	77 (55)	64 (58)	119 (48)	**4.17 (2.08-8.38)**	**0.001**	**3.63 (1.55-8.49)**	**0.003**	**2.38 (1.28-4-44)**	**0.006**	**2.76 (1.38-5.52)**	**0.004**
A/A	52 (37)	30 (28)	60 (24)	**5.59 (2.68-11.67)**	**0.001**	**4.16 (1.60-10.8)**	**0.003**	**2.21 (1.10-4.45)**	**0.036**	1.72 (0.76-3.87)	0.190
G/A+A/A	129 (92)	94 (85)	179 (72)	**4.65 (2.37-9.12)**	**0.001**	**3.83 (1.66-8.80)**	**0.002**	**2.33 (1.28-4.23)**	**0.004**	**2.18 (1.12-4.26)**	**0.022**

SNP rs7116432 Genotype	Colon n = 140 (%)	Rectum n = 110 (%)	Control n = 250 (%)	Colon vs. Control OR (95% CI)	P value	Colon vs. Control OR (95% CI)*	P value*	Rectum vs. Control OR (95% CI)	P value	Rectum vs Control OR (95% CI)*	P value*
A/A	39 (28)	32 (29)	102 (41)	1.00 (Reference)				1.00 (Reference)			
A/G	60 (43)	55 (50)	118 (47)	1.32 (0.82-2.15)	0.299	1.84 (0.96-3.50)	0.062	1.48 (0.89-2.47)	0.162	1.12 (0.63-1.99)	0.680
G/G	41 (29)	23 (21)	30 (12)	**3.57 (1.96-6.50)**	**0.001**	**5.94 (2.40-14.7)**	**0.001**	**2.22 (1.24-4.78)**	**0.013**	1.89 (0.86-4.16)	0.110
A/G+G/G	107 (72)	78 (71)	179 (72)	**1.78 (1.14-2.79)**	**0.014**	**2.46 (1.35-4.51)**	**0.003**	**2.17 (1.35-3.48)**	**0.001**	1.27 (0.74-2.16)	0.382

Bold values are statistically significant (P < 0.05).
*Adjust for age and sex

Table 6. Logistic regression analysis for the rs187116 and rs7116432 polymorphisms in the *CD44* gene.

Independent Variable	B[1]	s.e[2]	Wald[3]	d.f [4]	P value	OR (95% IC)
Colon vs. Rectum	0.543	0.263	4.270	1	**0.039**	**1.72 (1.02-2.88)**
Stage III+IV vs. I+II	–0.249	0.286	0.753	1	0.385	0.78 (0.44-1.36)
Age > 50 vs. < 50	0.315	0.333	0.898	1	0.343	1.37 (0.71-2.63)
rs187116						
GA+AA	0.357	0.414	0.741	1	0.390	1.42 (0.63-3.21)
Constant	–1.613	0.809	3.974	1	**0.046**	
Model	X^2 = 6.95 d.f.= 4 P = 0.138					
Colon vs. Rectum	0.517	0.264	3.829	1	**0.050**	1.67(1.00-2.81)
Stage III+IV vs. I+II	–0.266	0.288	0.855	1	0.355	0.76 (0.43-1.34)
Age > 50 vs. < 50	0.370	0.336	1.214	1	0.271	1.44 (0.74-2.80)
rs7116432						
AG+GG	0.466	0.289	2.603	1	0.107	1.59 (0.90-2.80)
Constant	–1.821	0.771	5.58	1	**0.018**	
Model	X^2 = 8.819 d.f.= 4 P = 0.066					

[1]regression coefficient.
[2]standard error.
[3]Wald Test.
[4]degrees of freedom.

tion is showed in Tables 4 and 5. In table 4, results adjustment by age and sex shown that patients in TNM I + II stages are associated with the G/A and A/A genotypes of the rs187116 polymorphism (OR = 4.12; 95% CI = 1.54-11.0, P = 0.005 and OR = 3.59; 95% CI = 1.25-10.3, P = 0.017, respectively); meanwhile, carriers of G/G genotype and under a dominant model of inheritance (A/G+G/G) of rs7116432 polymorphism were associated with advanced TNM stages (III + IV) (OR = > 3). Analysis adjustment by age and sex for tumor location showed that patients with colon localization of the tumor are associated with the G/A, and A/A genotypes for the rs187116 polymorphism (OR = 3.63; 95% CI = 1.55-8.49, P = 0.003 and OR = 4.16; 95% CI = 1.60-10.8, P = 0.003, respectively). These same results were observed for the patients with colon localization of the tumor and carrying the G/G genotype of the rs7116432 polymorphism (OR = 5.94; 95% CI = 2.40-14.7, P = 0.001).

Multivariate analysis of CD44 polymorphisms with clinicopathological characteristics in CRC patients

Table 6 shows the results of the multivariate analysis. Males patients with advanced TNM stages (III-IV) and older than 50 years that carry the G/A and A/A genotypes for the rs187116 polymorphism have a significantly increased risk (OR = 1.72; 95% CI = 1.02-2.88, P = 0.039). While, male patients carrying the genotypes A/G and G/G of the rs7116432 polymorphism are only marginally associated with advanced TNM stages in CRC (OR = 1.67; 95% CI = 1.00-2.81, P = 0.050).

CD44 haplotypes

Four different haplotypes in the *CD44* gene were observed (Table 7). Our results indicate that both loci rs187116 and rs7116432 are in linkage disequilibrium. The frequencies of the haplotypes were as follows A-A (CRC: 31.3%; controls: 31%), A-G (CRC: 29.7%; controls: 16.8%), G-A (CRC: 20.1%; controls: 33.4%), and G-G (CRC: 18.9%; controls: 18.8%). We observed that the combination of the A allele in rs187116 and the G allele in rs7116432 might be a risk factor for CRC susceptibility (OR = 2.08; 95% CI = 1.35-3.19, P = 0.000). However, the combination of the G allele in rs187116 and the A allele in rs7116432 have a protective effect against colorectal cancer (OR = 0.49; 95% CI = 0.32-0.74, P = 0.000).

Discussion

Currently, CSCs have been involved in the development, migration, invasion, resistance, aggressiveness and pluripotency of the malignant cells. *CD44* is a CSC gene with 50 kb long and 20 exons, located on the chromosome 11p13 (Basakran 2015); the CD44 protein in conjunction

Table 7. Association of the *CD44* gene haplotypes with colorectal cancer.

Haplotypes of the *CD44* gene		Frequencies		X^2	OR (95% CI)	P value
		CRC Group n = 250 (%)	**Control Group** n = 250 (%)			
rs187116- rs7116432						
A	**A**	78(31.3)	76 (31)	0.008	1.038 (0.71-1.51)	0.846
G	**A**	50 (20.1)	84 (33.4)	22.44	**0.494 (0.32-0.74)**	**0.001**
A	**G**	74 (29.7)	42 (16.8)	23.45	**2.08 (1.35-3.19)**	**0.001**
G	**G**	47 (18.9)	47 (18.8)	0.001	1.11 (0.70-1.77)	0.631

with their ligands (osteopontin and hyaluronan), interact with each other in a signaling network to trigger various tumorigenic processes as growth, differentiation, motility, survival, epithelial to mesenchymal transition (EMT) and metastasis (Zoller 2011; Zeilstra et al. 2013). In the intestinal mucosa, the CD44 protein is prominently expressed in stem cells and is a primary direct target of the Wnt signaling pathway (Wielenga et al. 1999; Zeilstra et al. 2008, 2013; van der Flier et al. 2009). Currently, there is accumulated evidence showing that CD44 is involved in the onset, progression and metastasis of intestinal tumors (Kim et al. 1994; Harada et al. 2001; Zeilstra et al. 2008, 2013; Todaro et al. 2010; Zoller 2011).

Several studies have demonstrated that *CD44* polymorphisms are significantly associated with susceptibility, prognosis and disease recurrence in different types of cancer. In CRC, the *CD44* rs187116 and rs7116432 polymorphisms have been analyzed only for prognosis and disease recurrence (Gerger et al. 2011; Stremitzer et al. 2015; Stotz et al. 2017), so that until our knowledge, the present study assesses, for the first time, the potential association of the *CD44* rs187116 and rs7116432 polymorphisms with susceptibility and clinicopathological parameters in Mexican patients with CRC.

From the 250 CRC patients here analyzed, 78.4% were ≥ 50 years old (average 58.6 years). Several studies have observed a high incidence of CRC in patients who were around 50 years old (Pourhoseingholi 2014; Siegel et al. 2014; Ferlay et al. 2015; Cancer Network Home of the Journal Oncology 2016; Gutierrez-Hurtado et al. 2016). The American Society of Clinical Oncology 2019 has stated that the average age at the time of diagnosis of colon cancer is 70 years, while, for rectal cancer, the average age is 63 years. As expected, the risk of colorectal cancer increases as people get to be older, although it can occur even in teenagers and young adults. In this study, a significant risk of cancer was observed in individuals over 50 years of age carrying the G/A and A/A genotypes of the rs187116 polymorphism; meanwhile, the patients carrying the G/G genotype in the rs7116432 polymorphism showed a significant CRC risk but in patients under 50. The increased susceptibility to cancer associated with patients younger than 50 years carrying the G/G genotype of the rs7116432 polymorphism could be related to a worse prognosis in these patients. It is possible to assume that the presence of this genotype would allow an earlier onset and a faster progression of the tumor, which would explain its more significant presence in patients younger than 50 years. On the other hand, the G/A, and A/A genotypes of the rs187116 polymorphism could induce a later and probably slower onset of CRC, thus explaining their preponderant presence in patients older than 50 years.

Regard gender, we found a slightly increased risk to develop CRC in males; this statistically significant risk regarding the control group was associated with the two polymorphisms analyzed in the *CD44* gene (rs187116 and rs7116432). These findings agree with the results reported by the American Society of Clinical Oncology (2019), accepting a marginally augmented risk of CRC in males. According to ESMO Clinical Practice Guidelines (Diagnosis, Treatment, and Follow-up), the estimated incidence of CRC is quite higher in males (ratio 4:1); furthermore, a 10-fold difference in incidence has been described for some regions (Labianca et al. 2013).

Our results shown that CRC patients in early TNM stages (I + II) more frequently carry the G/A and A/A genotypes of rs187116 polymorphism (OR = 4.12; 95% CI = 1.54-11.0, P = 0.001 and OR = 4.73; 95% CI= 1.81-12.3, P = 0.001 respectively); meanwhile, CRC patients in advanced TNM stages (III + IV), recurrently shown the G/G genotype of rs7116432 polymorphism (OR = 3.36; 95% CI = 1.51-7.49, P = 0.003). It is well known and accepted that CRC is a continuous multistep process that requires additional mutational changes and environmental factors that occur together over some time; therefore, it is not clear why, in the group of patients here studied, the early and advanced TNM stages exhibit different genotypes. Although we do not know when each patient included in this study began their malignant process, or how rapidly it evolved, it is possible to suppose that patients with CRC carrying the G/A and A/A genotypes of the rs187116 polymorphism could progress more slowly to through the first stages I and II before evolve on to advanced stages III or IV. Such a hypothesis could explain why the G/A, and A/A genotypes of rs187116 polymorphism are preponderantly observed in the early TNM stages. Although it is evident that the G/A and A/A genotypes of the rs187116 polymorphism represent a risk factor in developing CRC, the pres-

ence of these variants seems to provide some protection to the tumor cells, making them move more slowly from the early to the advanced stages. Conversely, CRC patients carrying the G/G genotype of the rs7116432 polymorphism could trigger biologic processes and induce mayor progress to reach stages III and IV in less time.

Interestingly, this situation coincides entirely with the results observed when comparing these genotypes with the age of patients with CRC. Thus: individuals who carry the G/A and A/A genotypes of rs187116 are over 50 years old with early TNM stages (I + II) (a less aggressive phenotype); meanwhile, individuals carrying the G/G genotype of the rs7116432 polymorphism are less than 50 years old and exhibit advanced TNM stage (III + IV) (a more aggressive phenotype).

For the first time, our results show that tobacco consumption is a risk factor for CRC in carriers of the A/A genotype of rs187116 polymorphism (OR = 4.80; 95% CI = 1.35-16.9, P = 0.022). Cigarette and alcohol consumption are potentially modifiable lifestyle factors that have been linked to colorectal cancer risk in multiple studies in Western populations (Tsong et al. 2007; Johnson et al. 2013; Cho et al. 2015). Moreover, Liang P. et al., in an extensive meta-analysis, found strong evidence that smoking is associated with an increased risk to CRC; however, no information demonstrates an association with the polymorphism rs187116 (Liang et al. 2009).

Regarding the location of the tumor, the presence of G/A and A/A genotypes for the rs187116 polymorphism (OR = 3.63; 95% CI = 1.55-8.49, P = 0.003, OR = 4.16; 95% CI = 1.60-10.8, P = 0.003) respectively and the presence of G/G genotype for rs7116432 polymorphism (OR = 5.94; 95% CI = 2.40-14.7, P = 0.001) were associated with an increased risk for tumor development in the colon. In support of these results, studies achieved in Western countries found that two-thirds of CRC is in the colon, and only one third in the rectum (Paschke et al. 2018). Such a difference could be explained with the considerable length of the colon and consequently, a more extensive mucosa where a tumor could eventually start. It is also clear that, from a biological and histopathological point of view, the colon and rectum are distinct entities (Slattery et al. 2011). In previous studies realized in the Mexican population, we have observed that patients with CRC also showed a greater susceptibility to develop tumors in the colon. (Rosales-Reynoso et al. 2019 a, b).

The multivariate analysis performed in our study showed that male patients carrying risk genotypes for the rs187116 and rs7116432 polymorphisms have a significant risk of developing colon cancer specifically (OR = 1.72; 95% CI = 1.02-2.88, P = 0.039; OR = 1.67; 95% CI = 0.99-2.81, P = 0.050) respectively.

The haplotype analysis showed that the combination of risk alleles (haplotype A-G) of the rs187116 and rs7116432 polymorphisms is associated with an increased risk to CRC (OR = 2.08; 95% CI = 1.35-3.19, P = 0.000); while the combination of wild-type alleles (G-A haplotype) is associated with protection.

For the first time, an association of the *CD44* rs187116 and rs7116432 polymorphisms with CRC risk and with some clinic-pathological features of CRC is demonstrated in this study. These same polymorphisms have also been associated with an increased risk of breast and gastric cancer, as reported by several studies in different populations (Winder et al. 2011; Suenaga et al. 2015; Tongtawee et al. 2017; Sapcharoen et al. 2019).

In conclusion, the analysis of results in this study reveals that the *CD44* rs187116 and rs7116432 polymorphisms mean genetic risk factors for CRC. Some of these genotypes are also associated with the tumor location and with the TNM stage in these patients. Although additional studies including larger samples and functional analysis of the polymorphisms are necessary to confirm and extend our findings, it is reasonable to propose that the *CD44* rs187116 and rs7116432 SNPs could be considered useful biomarkers of prognosis and tumor location in CRC. A limitation of this study was the absence of follow-up data and treatment response outcomes in these patients.

Acknowledgments

This study was supported by Grants from "Fondo de Investigación en Salud del Instituto Mexicano del Seguro Social" (FIS/IMSS/PROT/G18/1822).

Conflict of Interest

The authors declare no conflict of interest.

References

American Society of Clinical Oncology (2019) Guidelines by clinical area, gastrointestinal cancer. https://www.asco.org/research-guidelines/quality-guidelines/guidelines [*Accessed:* August 24, 2019].

Atashzar, M.R., Baharlou, R., Karami, J., Abdollahi, H., Rezaei, R., Pourramezan, F. & Zoljalali Moghaddam, S.H. (2020) Cancer stem cells: a review from origin to therapeutic implications. *J. Cell. Physiol.*, **235**, 790-803.

Basakran, N.S. (2015) CD44 as a potential diagnostic tumor marker. *Saudi Med. J.*, **36**, 273-279.

Bitaraf, S.M., Mahmoudian, R.A., Abbaszadegan, M., Mohseni Meybodi, A., Taghehchian, N., Mansouri, A., Forghanifard, M.M., Memar, B. & Gholamin, M. (2018) Association of two CD44 polymorphisms with clinical outcomes of gastric cancer patients. *Asian Pac. J. Cancer Prev.*, **19**, 1313-1318.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. & Jemal, A. (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, **68**, 394-424.

Cancer Network Home of the Journal Oncology (2016) Colon, rectal, and anal cancers. Cancer management. http://www.cancernetwork.com/cancer-management/colon-rectal-and-anal-cancers/ [*Accessed:* August 11, 2019].

Cho, S., Shin, A., Park, S.K., Shin, H.R., Chang, S.H. & Yoo, K.Y. (2015) Alcohol drinking, cigarette smoking and risk of colorectal cancer in the Korean multi-center cancer cohort. *J. Cancer Prev.*, **20**, 147-152.

Chou, Y.E., Hsieh, M.J., Chiou, H.L., Lee, H.L., Yang, S.F. &

Chen, T.Y. (2014) CD44 gene polymorphisms on hepatocellular carcinoma susceptibility and clinicopathologic features. *Biomed. Res. Int.*, **2014**, 231474.

Dalerba, P., Dylla, S.J., Park, I.K., Liu, R., Wang, X., Cho, R.W., Hoey, T., Gurney, A., Huang, E.H., Simeone, D.M., Shelton, A.A., Parmiani, G., Castelli, C. & Clarke, M.F. (2007) Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl. Acad. Sci. USA*, **104**, 10158-10163.

Fearon, E.R. & Vogelstein, B. (1990) A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759-767.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D. & Bray, F. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer*, **136**, E359-386.

Fischer, J., Walker, L.C., Robinson, B.A., Frizelle, F.A., Church, J.M. & Eglinton, T.W. (2019) Clinical implications of the genetics of sporadic colorectal cancer. *ANZ J. Surg.*, **89**, 1224-1229.

Freedman, J.A., Wang, Y., Li, X., Liu, H., Moorman, P.G., George, D.J., Lee, N.H., Hyslop, T., Wei, Q. & Patierno, S.R. (2018) Single-nucleotide polymorphisms of stemness genes predicted to regulate RNA splicing, microRNA and oncogenic signaling are associated with prostate cancer survival. *Carcinogenesis*, **39**, 879-888.

Gerger, A., Zhang, W., Yang, D., Bohanes, P., Ning, Y., Winder, T., LaBonte, M.J., Wilson, P.M., Benhaim, L., Paez, D., El-Khoueiry, R., El-Khoueiry, A., Kahn, M. & Lenz, H.J. (2011) Common cancer stem cell gene variants predict colon cancer recurrence. *Clin. Cancer Res.*, **17**, 6934-6943.

Gutierrez-Hurtado, I.A., Puebla-Perez, A.M., Delgado-Saucedo, J.I., Figuera, L.E., Zuniga-Gonzalez, G.M., Gomez-Mariscal, K., Ronquillo-Carreon, C.A. & Gallegos-Arreola, M.P. (2016) Association between TNF-alpha-308G>A and -238G>A gene polymorphisms and TNF-alpha serum levels in Mexican colorectal cancer patients. *Genet. Mol. Res.*, **15**, gmr8199.

Han, S., Yang, W., Zong, S., Li, H., Liu, S., Li, W., Shi, Q. & Hou, F. (2017) Clinicopathological, prognostic and predictive value of CD166 expression in colorectal cancer: a meta-analysis. *Oncotarget*, **8**, 64373-64384.

Harada, N., Mizoi, T., Kinouchi, M., Hoshi, K., Ishii, S., Shiiba, K., Sasaki, I. & Matsuno, S. (2001) Introduction of antisense CD44S CDNA down-regulates expression of overall CD44 isoforms and inhibits tumor growth and metastasis in highly metastatic colon carcinoma cells. *Int. J. Cancer*, **91**, 67-75.

Hen, O. & Barkan, D. (2019) Dormant disseminated tumor cells and cancer stem/progenitor-like cells: Similarities and opportunities. *Semin. Cancer Biol.*, doi: 10.1016/j.semcancer.2019.09.002 [Epub ahead of print].

Huh, J.W., Kim, H.R., Kim, Y.J., Lee, J.H., Park, Y.S., Cho, S.H. & Joo, J.K. (2009) Expression of standard CD44 in human colorectal carcinoma: association with prognosis. *Pathol. Int.*, **59**, 241-246.

Johnson, C.M., Wei, C., Ensor, J.E., Smolenski, D.J., Amos, C.I., Levin, B. & Berry, D.A. (2013) Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control*, **24**, 1207-1222.

Kim, H., Yang, X.L., Rosada, C., Hamilton, S.R. & August, J.T. (1994) CD44 expression in colorectal adenomas is an early event occurring prior to K-ras and p53 gene mutation. *Arch. Biochem. Biophys.*, **310**, 504-507.

Labianca, R., Nordlinger, B., Beretta, G.D., Mosconi, S., Mandala, M., Cervantes, A., Arnold, D. & Group, E.G.W. (2013) Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.*, **24** Suppl 6, vi64-72.

Liang, P.S., Chen, T.Y. & Giovannucci, E. (2009) Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. *Int. J. Cancer*, **124**, 2406-2415.

Lin, V.C., Huang, S.P., Huang, C.Y., Yu, C.C., Yin, H.L., Huang, T.Y., Lee, C.H., Lu, T.L. & Bao, B.Y. (2017) Cancer stem cell gene variants predict disease recurrence in patients treated with radical prostatectomy for prostate cancer. *Int. J. Med. Sci.*, **14**, 1301-1306.

Liu, Y., Qing, H., Su, X., Wang, C., Li, Z. & Liu, S. (2015) Association of CD44 gene polymorphism with survival of NSCLC and risk of bone metastasis. *Med. Sci. Monit.*, **21**, 2694-2700.

Miller, S.A., Dykes, D.D. & Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.*, **16**, 1215.

Mueller, M.T., Hermann, P.C., Witthauer, J., Rubio-Viqueira, B., Leicht, S.F., Huber, S., Ellwart, J.W., Mustafa, M., Bartenstein, P., D'Haese, J.G., Schoenberg, M.H., Berger, F., Jauch, K.W., Hidalgo, M. & Heeschen, C. (2009) Combined targeted treatment to eliminate tumorigenic cancer stem cells in human pancreatic cancer. *Gastroenterology*, **137**, 1102-1113.

Orian-Rousseau, V. & Schmitt, M. (2015) CD44 regulates Wnt signaling at the level of LRP6. *Mol. Cell. Oncol.*, **2**, e995046.

Paschke, S., Jafarov, S., Staib, L., Kreuser, E.D., Maulbecker-Armstrong, C., Roitman, M., Holm, T., Harris, C.C., Link, K.H. & Kornmann, M. (2018) Are colon and rectal cancer two different tumor entities? A proposal to abandon the term colorectal cancer. *Int. J. Mol. Sci.*, **19**, E2577.

Pourhoseingholi, M.A. (2014) Epidemiology and burden of colorectal cancer in Asia-Pacific region: what shall we do now? *Transl. Gastrointest. Cancer*, **3**, 169-173.

Reya, T., Morrison, S.J., Clarke, M.F. & Weissman, I.L. (2001) Stem cells, cancer, and cancer stem cells. *Nature*, **414**, 105-111.

Ricci-Vitiani, L., Lombardi, D.G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C. & De Maria, R. (2007) Identification and expansion of human colon-cancer-initiating cells. *Nature*, **445**, 111-115.

Rosales-Reynoso, M.A., Saucedo-Sarinana, A.M., Contreras-Diaz, K.B., Marquez-Gonzalez, R.M., Barros-Nunez, P., Pineda-Razo, T.D., Marin-Contreras, M.E., Duran-Anguiano, O., Gallegos-Arreola, M.P., Flores-Martinez, S.E. & Sanchez-Corona, J. (2019a) Genetic polymorphisms in APC, DVL2, and AXIN1 are associated with susceptibility, advanced TNM stage or tumor location in colorectal cancer. *Tohoku J. Exp. Med.*, **249**, 173-183.

Rosales-Reynoso, M.A., Zepeda-Lopez, P., Saucedo-Sarinana, A.M., Pineda-Razo, T.D., Barros-Nunez, P., Gallegos-Arreola, M.P., Flores-Martinez, S.E. & Sanchez-Corona, J. (2019b) GSK3beta polymorphisms are associated with tumor site and TNM stage in colorectal cancer. *Arch. Iran. Med.*, **22**, 453-460.

Rossari, F., Zucchinetti, C., Buda, G. & Orciuolo, E. (2019) Tumor dormancy as an alternative step in the development of chemoresistance and metastasis - clinical implications. *Cell. Oncol.* (*Dordr.*), doi.org/10.1007/s13402-019-00467-7.

Sanders, M.A. & Majumdar, A.P. (2011) Colon cancer stem cells: implications in carcinogenesis. *Front. Biosci.* (*Landmark Ed*), **16**, 1651-1662.

Sapcharoen, K., Sanguansermsri, P., Yasothornsrikul, S., Muisuk, K. & Srikummool, M. (2019) Gene combination of CD44 rs187116, CD133 rs2240688, NF-kappaB1 rs28362491 and GSTM1 deletion as a potential biomarker in risk prediction of breast cancer in lower Northern Thailand. *Asian Pac. J. Cancer Prev.*, **20**, 2493-2502.

Siegel, R., Desantis, C. & Jemal, A. (2014) Colorectal cancer statistics, 2014. *CA Cancer J. Clin.*, **64**, 104-117.

Skandalis, S.S., Karalis, T.T., Chatzopoulos, A. & Karamanos, N.K. (2019) Hyaluronan-CD44 axis orchestrates cancer stem cell functions. *Cell. Signal.*, **63**, 109377.

Slattery, M.L., Wolff, E., Hoffman, M.D., Pellatt, D.F., Milash, B. & Wolff, R.K. (2011) MicroRNAs and colon and rectal cancer: differential expression by tumor location and subtype.

Genes Chromosomes Cancer, **50**, 196-206.

Spring, F.A., Dalchau, R., Daniels, G.L., Mallinson, G., Judson, P.A., Parsons, S.F., Fabre, J.W. & Anstee, D.J. (1988) The Ina and Inb blood group antigens are located on a glycoprotein of 80,000 MW (the CDw44 glycoprotein) whose expression is influenced by the In(Lu) gene. *Immunology*, **64**, 37-43.

Stotz, M., Herzog, S.A., Pichler, M., Smolle, M., Riedl, J., Rossmann, C., Bezan, A., Stoger, H., Renner, W., Berghold, A. & Gerger, A. (2017) Cancer stem cell gene variants in CD44 predict outcome in Stage II and Stage III colon cancer patients. *Anticancer Res.*, **37**, 2011-2018.

Stremitzer, S., Zhang, W., Yang, D., Ning, Y., Stintzing, S., Sunakawa, Y., Sebio, A., Yamauchi, S., Matsusaka, S., Parekh, A., Barzi, A., El-Khoueiry, R., Stift, J., Wrba, F., Gruenberger, T., et al. (2015) Variations in genes involved in dormancy associated with outcome in patients with resected colorectal liver metastases. *Ann. Oncol.*, **26**, 1728-1733.

Suenaga, M., Yamada, S., Fuchs, B.C., Fujii, T., Kanda, M., Tanaka, C., Kobayashi, D., Fujiwara, M., Tanabe, K.K. & Kodera, Y. (2015) CD44 single nucleotide polymorphism and isoform switching may predict gastric cancer recurrence. *J. Surg. Oncol.*, **112**, 622-628.

Todaro, M., Francipane, M.G., Medema, J.P. & Stassi, G. (2010) Colon cancer stem cells: promise of targeted therapy. *Gastroenterology*, **138**, 2151-2162.

Tongtawee, T., Wattanawongdon, W., Simawaranon, T., Kaewpitoon, S., Kaengpenkae, S., Jintabanditwong, N., Tangjanyatham, P., Ratchapol, W., Kangwantas, K., Dechsukhum, C., Leeanansaksiri, W., Kaewpitoon, N., Matrakool, L. & Panpimanmas, S. (2017) Expression of cancer stem cell marker CD44 and its polymorphisms in patients with chronic gastritis, precancerous gastric lesion, and gastric cancer: a cross-sectional multicenter study in Thailand. *Biomed. Res. Int.*, **2017**, 4384823.

Tsong, W.H., Koh, W.P., Yuan, J.M., Wang, R., Sun, C.L. & Yu, M.C. (2007) Cigarettes and alcohol in relation to colorectal cancer: the Singapore Chinese Health Study. *Br. J. Cancer*, **96**, 821-827.

Tulsyan, S., Agarwal, G., Lal, P., Agrawal, S., Mittal, R.D. & Mittal, B. (2013) CD44 gene polymorphisms in breast cancer risk and prognosis: a study in North Indian population. *PLoS One*, **8**, e71073.

van der Flier, L.G., van Gijn, M.E., Hatzis, P., Kujala, P., Haegebarth, A., Stange, D.E., Begthel, H., van den Born, M., Guryev, V., Oving, I., van Es, J.H., Barker, N., Peters, P.J., van de Wetering, M. & Clevers, H. (2009) Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell*, **136**, 903-912.

Vermeulen, L., Todaro, M., de Sousa Mello, F., Sprick, M.R., Kemper, K., Perez Alea, M., Richel, D.J., Stassi, G. & Medema, J.P. (2008) Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. *Proc. Natl. Acad. Sci. USA*, **105**, 13427-13432.

Wahab, S.M.R., Islam, F., Gopalan, V. & Lam, A.K. (2017) The identifications and clinical implications of cancer stem cells in colorectal cancer. *Clin. Colorectal Cancer*, **16**, 93-102.

Wan, Q., Zhang, D., Zhou, Q., Li, M., Wang, Y., Song, Y. & Xu, T. (2019) Association of CD44 gene rs187115 polymorphism with colorectal cancer risk and prognosis in Chinese Han population: a case-control study. *Aging (Albany NY)*, **11**, 9616-9625.

Wielenga, V.J., Smits, R., Korinek, V., Smit, L., Kielman, M., Fodde, R., Clevers, H. & Pals, S.T. (1999) Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. *Am. J. Pathol.*, **154**, 515-523.

Winder, T., Ning, Y., Yang, D., Zhang, W., Power, D.G., Bohanes, P., Gerger, A., Wilson, P.M., Lurje, G., Tang, L.H., Shah, M. & Lenz, H.J. (2011) Germline polymorphisms in genes involved in the CD44 signaling pathway are associated with clinical outcome in localized gastric adenocarcinoma. *Int. J. Cancer*, **129**, 1096-1104.

World Health Organization (2018) Estimated age-standardized incidence and mortality rates (World) in 2018, worldwide, males, all ages.
https://gco.iarc.fr/today/data/factsheets/populations/900-world-fact-sheets.pdf
[*Accessed:* October 24, 2019].

Zeilstra, J., Joosten, S.P., Dokter, M., Verwiel, E., Spaargaren, M. & Pals, S.T. (2008) Deletion of the WNT target and cancer stem cell marker CD44 in Apc(Min/+) mice attenuates intestinal tumorigenesis. *Cancer Res.*, **68**, 3655-3661.

Zeilstra, J., Joosten, S.P., Vermeulen, L., Koster, J., Medema, J.P., Versteeg, R., Spaargaren, M. & Pals, S.T. (2013) CD44 expression in intestinal epithelium and colorectal cancer is independent of p53 status. *PLoS One*, **8**, e72849.

Zhou, J., Nagarkatti, P.S., Zhong, Y., Zhang, J. & Nagarkatti, M. (2011) Implications of single nucleotide polymorphisms in CD44 exon 2 for risk of breast cancer. *Eur. J. Cancer Prev.*, **20**, 396-402.

Zoller, M. (2011) CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? *Nat. Rev. Cancer*, **11**, 254-267.