RESEARCH ARTICLE

Calpain-10 SNP43 and SNP19 Polymorphisms and Colorectal Cancer: a Matched Case-control Study

Xiao-Qin Hu[1], Ping Yuan[1], Rong-Sheng Luan[1], Xiao-Ling Li[1], Wen-Hui Liu[1], Fei Feng[1], Jin Yan[2], Yan-Fang Yang[1]*

Abstract

Objective: Insulin resistance (IR) is an established risk factor for colorectal cancer (CRC). Given that CRC and IR physiologically overlap and the calpain-10 gene (CAPN10) is a candidate for IR, we explored the association between CAPN10 and CRC risk. Methods: Blood samples of 400 case-control pairs were genotyped, and the lifestyle and dietary habits of these pairs were recorded and collected. Unconditional logistic regression (LR) was used to assess the effects of CAPN10 SNP43 and SNP19, and environmental factors. Both generalized multifactor dimensionality reduction (GMDR) and the classification and regression tree (CART) were used to test gene-environment interactions for CRC risk. Results: The GA+AA genotype of SNP43 and the Del/Ins+Ins/Ins genotype of SNP19 were marginally related to CRC risk (GA+AA: OR = 1.35, 95% CI = 0.92-1.99; Del/Ins+Ins/Ins: OR = 1.31, 95% CI = 0.84-2.04). Notably, a high-order interaction was consistently identified by GMDR and CART analyses. In GMDR, the four-factor interaction model of SNP43, SNP19, red meat consumption, and smoked meat consumption was the best model, with a maximum cross-validation consistency of 10/10 and testing balance accuracy of 0.61 ($P < 0.01$). In LR, subjects with high red and smoked meat consumption and two risk genotypes had a 6.17-fold CRC risk (95% CI = 2.44-15.6) relative to that of subjects with low red and smoked meat consumption and null risk genotypes. In CART, individuals with high smoked and red meat consumption, SNP19 Del/Ins+Ins/Ins, and SNP43 GA+AA had higher CRC risk (OR = 4.56, 95%CI = 1.94-10.75) than those with low smoked and red meat consumption. Conclusions: Though the single loci of CAPN10 SNP43 and SNP19 are not enough to significantly increase the CRC susceptibility, the combination of SNP43, SNP19, red meat consumption, and smoked meat consumption is associated with elevated risk.

Keywords: Calpain-10 gene - genetic polymorphism - interaction - colorectal cancer

Asian Pac J Cancer Prev, **14 (11)**, 6673-6680

Introduction

As the second most common cancer and fourth leading cause of cancer deaths in the world (Ferlay et al., 2010), colorectal cancer (CRC) is an important health issue. More than half of CRC cases occur in developed countries, but the incidence of CRC in developing countries is also steadily rising. In China, the incidence of CRC grew from $15.0/10^5$ to $32.5/10^5$ among men and from $9.7/10^5$ to $26.7/10^5$ among women from 2005 to 2007. Accordingly, mortality rates increased from $8.6/10^5$ to $15.6/10^5$ among men and from $5.4/10^5$ to $12.7/10^5$ among women (Zhao and Zhang, 2011).

Individuals with obesity and type 2 diabetes mellitus (T2DM) have high incidence and mortality rates of CRC compared with health people (Basen-Engquist and Chang, 2011; Noto et al., 2011), whereas both obesity and T2DM are associated with insulin resistance (IR) (Wilkin, 2009; Zeyda and Stulnig, 2009). In the 1990s, McKeown-Eyssen (1994) and Giovannucci (1995) originally proposed a hypothesis of "insulin resistance-colon cancer". IR, which is characterized by compensatory hyperinsulinemia, is believed to promote carcinogenesis through several potential mechanisms (Djiogue et al., 2013). First, hyperinsulinemia and a high insulin-like growth factor-1 (IGF-1) level, which are caused by increased insulin, promote cell proliferation and inhibit apoptosis through activating the protein kinase B (Akt) and mitogen-activated protein kinase signaling networks in neoplastic tissues (Pollak, 2008; Leng et al., 2001). Second, IR is associated with inhibition of the liver kinase B1 (LKB1)-adenosine 5'-monophosphate (AMP)-activated protein kinase (LKB1-AMPK) pathway which promotes energy storage and obesity. In experimental studies, inhibition of the LKB1-AMPK was associated with protein synthesis and cancer development (Kuhajda, 2008; Wang and Guan, 2009). Third, the inflammatory environment of patients with obesity and T2DM can promote carcinogenesis (Tuncman et al., 2006). Many studies elaborate on the relationship between IR and CRC risk. A study on

*[1]Department of Epidemiology, West China School of Public Health, Sichuan University, [2]Department of Intestinal Surgery, Sichuan Cancer Hospital, Chengdu, Sichuan, China *For correspondence: yang2009@scu.edu.cn*

gastrin gene knockout (GAS-KO) mice indicated that hyperinsulinemia, as a result of the amidated gastrins loss, may increase colon carcinogenesis (Cowey et al., 2005). A recent study suggested that central obesity and IR played important roles in the early stage of colorectal neoplasia, particularly among men (Ortiz et al., 2012). Komninou et al. (2003) performed a review on the contribution of IR to colon carcinogenesis and provided strong support for the hypothesis.

Based on this hypothesis mentioned above, CRC and IR may share genetic susceptibility factors associated with both pathologies. The calpain-10 (CAPN10) gene is the first candidate T2DM gene identified through genome-wide linkage and positional cloning (Horikawa et al., 2000). This gene is related to impaired glucose tolerance and IR phenotypes in normal individuals (Turner et al., 2005; Sáez et al., 2008). CAPN10 is cloned at the non-insulin-dependent diabetes mellitus 1 (NIDDM1) region within D2S125 to D2S140 and is thus suspected to be a susceptibility gene for IR (Hanis et al., 1996). After doing extensive article review, we chose 2 of the most studied hotspot loci on CAPN10 to investigate the association between gene polymorphisms of CAPN10 and the susceptibility to CRC. These single nucleotide polymorphisms (SNP) are CAPN10 SNP43 (rs3792267) and SNP19 (rs3842570), the independent impacts of which on IR were putative.

Environmental and genetic factors place certain individuals at higher risk of developing CRC (Kotnis et al., 2005). The World Cancer Research Fund asserted that red meat consumption increases the risk of CRC (WCRF, 2007). High red meat consumption is also considered a potential risk factor for IR (de Oliveira Otto et al., 2012; Lajous et al., 2012). Smoked meat (meat salted for several days and then smoked with burning hardwood) is an integral part of the diet of Sichuan families. However, smoked meat is related to the risk of CRC (Santarelli et al., 2010; Santarelli et al., 2013). Therefore, red and smoked meat consumption and CAPN10 polymorphisms were used in this study to explore the gene-environment interaction on CRC.

To our knowledge, this is the first study reporting the association of gene polymorphisms of the IR pathway with CRC risk. The gene-gene and gene-environment interaction is a hot topic in genetic epidemiology, thus, we examined matched case-control pairs to elucidate the relationship between CAPN10 and CRC and determine the possible interaction among SNP43, SNP19, red meat consumption, and smoked meat consumption.

Materials and Methods

Study Population

Four hundred cases of CRC and 400 age-matched (±3 years) and gender-matched controls were invited to participate in the study. All the subjects were Han Chinese aged 20-80 and had been living in Sichuan for at least 20 years. From July 2010 to May 2012, cases with newly histopathologically diagnosed primary CRC were consecutively selected from the Sichuan Cancer Hospital (Chengdu, Sichuan, China). The controls were cancer-free individuals selected from a group of healthy people receiving routine medical examinations at the Zhonghe Community Health Service Center (Chengdu, Sichuan, China) during the same period. Subjects with a history of cancer were excluded from the study.

All subjects provided written informed consent before completing the questionnaire survey and laboratory tests. The study protocol was approved by the Institutional Review Board of Sichuan University.

Exposure to environmental factors

Professionally trained interviewers used a structured questionnaire to interview each subject face-to-face and followed a written protocol to guide ascertainment and reduce surveillance, interviewer, and recall bias. The questionnaire primarily included questions on demographic factors and potential CRC risk factors, such as family history of CRC (i.e., among first- and second-degree relatives); sitting hours a day; smoking, alcohol-drinking, and tea-drinking habits; and red meat (e.g., beef, lamb, and pork) and smoked meat consumption. Smoking, alcohol-drinking, and tea-drinking habits were measured according to duration, type, and consumption. Smoking was defined as smoking more than one cigarette a day for at least six months, drinking as drinking alcohol more than twice a week for at least six months, and tea drinking as drinking tea more than once a day for at least six months, currently or before. Consumption of red and smoked meat was measured by frequency a week; one intake of such meat was defined as more than 50 and 25 g, respectively, according to the dietary habits of Sichuan residents. Considering the median of red meat consumption in the controls and the 50 g to 75 g daily red meat consumption recommended by the Dietary Guideline and Balance Diet Pagoda for Chinese Residents, red meat consumption was divided into a dichotomized variable with seven times a week as the cutoff point. Individuals who consumed red meat more than seven times a week were categorized as the high red meat consumption group; those with a lower frequency were considered the low red meat consumption group. Sichuan residents often consume significant amounts of smoked meat in the months before and after the Spring Festival. Thus, smoked meat consumption was measured according to the number of consumption months a year and the frequency a week during these months. Considering the medians of smoked meat consumption in the controls, subjects who consumed equal to or more than two months a year and equal to or more than three times a week during these months were categorized as the high smoked meat consumption group; those with a lower frequency were categorized as the low smoked meat consumption group. All subjects were asked about their lifestyle and habits relative to the same 10 years before disease diagnosis.

DNA extraction and genotyping

Whole blood samples (2 mL) were drawn from each participant via venipuncture into a tube containing trisodium citrate and stored at –40 °C. Genomic DNA was then extracted with an SE blood DNA kit (Tiangen Biotech, Beijing) according to manufacturer's instructions.

The polymerase chain reaction-restriction fragment length polymorphism (PCR–RFLP) method was used to detect the SNP43 and SNP19 genotypes. The intron 3 region containing SNP43 was amplified with forward primer 5'-GCTGGCTGGTGACATCAGTGC-3' and reverse primer 5'-ACCAAGTCAAGGCTTAGCCTCACCTTCA TA-3' (product size = 254 bp). Reactions were performed at a final volume of 25 μl, with 2 μl genomic DNA (0.2 μg), 12.5 μl 2×Taq MasterMix (Tiangen Biotech, Beijing), 8.5 μl RNase free water, and 1 μl of each primer (10 μM). The cycling conditions were 94 °C for 5 min; 32 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s; and 72 °C for 10 min. The PCR products were digested with 5 U NdeI (NEB, Beijing) at 37 °C for 16 h. The digested products were separated on a 2% agarose gel. The G allele had a size of 254 bp, and the A allele had 223 and 31 bp. The intron 6 region containing SNP19 was amplified with forward primer 5'-GTTTGGTTCTCTTCAGCGTGGAG -3' and reverse primer 5'-CATGAACCCTGGCAGGGT CTAAG-3' (product size = 187 bp). The reaction system was the same as that of SNP43. The cycling conditions were 94 °C for 12 min; 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s; and 72 °C for 10 min. The PCR products were separated on a 2% agarose gel. The Del allele (two repeats of a 32-bp sequence) had a size of 155 bp, and the Ins allele (three repeats) had 187 bp.

A total of 160 random samples (20% of the total subjects) were sequenced by an ABI 3730XL sequencer (Applied Biosystems, Invitrogen Trading, Shanghai, China) to confirm the accuracy of the genotype. The concordance rate was 100%.

Statistical analysis

The Hardy-Weinberg equilibrium (HWE) of the genotype distributions of SNP43 and SNP19 in the control group was assessed by a χ^2 goodness-of-fit test through online software SHEsis (http://analysis.Bio-x.cn/myAnalysis.php). Haploview 4.2 (http://www.broadinstitute.org/haploview/haploview-downloads) was used to estimate linkage disequilibrium (LD) between SNP43 and SNP19. With the wild genotype as reference, the genotype was analyzed as a categorical variable and then reanalyzed as a dichotomized variable by combining the heterozygous and homozygous genotypes of the variants. The marginal effects of environmental factors and genotypes on CRC risk were estimated by unconditional logistic regression (LR), and their ORs with 95% CIs were adjusted according to family per capita annual income, family history of CRC, sitting hours per day, body mass index (BMI), and smoking, alcohol-drinking, and tea-drinking habits. LR analysis was performed with SPSS 18.0 (SPSS, Chicago, IL). The significance level in all tests was $P < 0.05$.

To estimate high-order gene-environment interactions, we used generalized multifactor dimensionality reduction (GMDR 0.7, http://www.healthsystem.virginia.edu/internet/addiction-genomics/software/), which is described in detail elsewhere (Lou et al., 2007). GMDR reduces high-dimensional data to a single dimensional variable with two levels (high risk and low risk) by the ratio of patients to controls and thereby facilitates the detection of interactions in small sample sizes. The GMDR software provides several parameters, including testing balance accuracy (TBA), sign test *P* value, and cross-validation consistency (CVC) to assess each selected interaction model. The model with the maximum TBA, a sign test *P* value of <0.05, and a CVC of 10/10 was considered the best model. The above confounding factors were included as covariates in GMDR analysis for gene-environment interactions. The combined effect of the variables in the best model based on the number of risk genotypes modified by environmental factors was also evaluated through LR analysis.

Classification and regression tree (CART) analysis was performed with Salford Predictive Modeler 7 (http://www.salford-systems.com/en/products/cart) to explore possible high-order gene-environment interactions. CART is a binary recursive partitioning method that produces a decision tree to identify specific combinations of genes and environmental factors associated with CRC risk. The decision tree was split until terminal nodes had no subsequent statistically significant splits or until these nodes reached a pre-specified minimum size of 10 subjects. The data were randomly divided into a learning part (90%) and a testing part (10%). The learning part was used to construct the tree model, and the testing part was used to internally validate the tree model. Individuals with different CRC risks were identified in different terminal nodes of the tree, indicating that potential interactions may exist among genes and environmental factors. The risk of each node was also evaluated by LR analysis. ORs and 95% CIs were adjusted according to the confounding factors, with the node accounting for the least percentage of cases used as the reference.

Results

Characteristics of participants

Given that the cases and controls were matched in terms of age and sex, no substantial differences in these variables were found between the cases and controls. Among the participants, 233 (58.2%) were males and 167 (41.8%) were females. The mean ages were 55.7±11.1 years for the cases and 55.7±11.2 years for the controls ($t = 0.01$, $P = 0.99$). The cases included 268 (67.0%) subjects with rectal cancer and 132 (33.0%) subjects with colon cancer. The characteristics of the participants are shown in Table 1. Having a family history of CRC and long sitting hours a day (≥8 h/d) were associated with increased CRC risk (CRC history: OR = 3.78, 95% CI = 1.76-8.15; sitting hours: OR = 1.97, 95% CI = 1.35-2.88). By contrast, habitual tea drinking decreased CRC risk (OR = 0.60, 95% CI = 0.43-0.82). People consuming red meat more than seven times a week had elevated CRC risk (OR = 1.78, 95% CI = 1.32-2.40) compared with those consuming red meat at a lower frequency. Compared with the participants with low smoked meat consumption, those with high smoked meat consumption had a 2.35-fold increase (95% CI = 1.71-3.23) in CRC risk. No statistically significant association was found between CRC risk and family per capita annual income, BMI, smoking habit, and alcohol-drinking habit.

Table 1. ORs and 95%CIs of Main Risk Factors for CRC

Variable	Cases N(%)	Controls N(%)	OR (95%CI)	*P*
Family per capita annual income (yuan/year)				
<10000	134(33.5)	118(29.5)	Reference	0.16
≥10000	266(66.5)	282(70.5)	0.79(0.58-1.09)	
Family history of CRC				
No	365(91.2)	391(97.8)	Reference	<0.01
Yes	35(8.8)	9(2.2)	3.78(1.76-8.15)	
Sitting hours per day				
<8	296(74.0)	337(84.2)	Reference	<0.01
≥8	104(26.0)	63(15.8)	1.97(1.35-2.88)	
BMI				
<24	245(61.2)	231(57.8)	Reference	0.33
≥24	155(38.8)	169(42.2)	0.86(0.64-1.16)	
Smoking habit				
No	226(56.5)	230(57.5)	Reference	0.46
Yes	174(43.5)	170(42.5)	1.15(0.80-1.65)	
Alcohol-drinking habit				
No	262(65.5)	259(64.8)	Reference	0.50
Yes	138(34.5)	141(35.2)	0.88(0.61-1.28)	
Tea-drinking habit				
No	228(57.0)	190(47.5)	Reference	<0.01
Yes	172(43.0)	210(52.5)	0.60(0.43-0.82)	
Red meat				
Low	144(36.0)	214(53.5)	Reference	<0.01
High	256(64.0)	186(46.5)	1.78(1.32-2.40)	
Smoked meat				
Low	232(58.0)	303(75.8)	Reference	<0.01
High	168(42.0)	97(24.2)	2.35(1.71-3.23)	

Genotypes

Table 2 shows the associations between the genotypes and CRC. The genotype distributions among the controls were in HWE for both polymorphisms (SNP43: $\chi^2 = 0.81$, $P = 0.37$; SNP19: $\chi^2 = 0.10$, $P = 0.75$). The r^2 and D' of LD between SNP43 and SNP19 were 0.05 and 0.99, respectively. Unlike of the GA genotype (OR = 1.25, 95% CI = 0.83-1.86), the OR of the AA genotype of SNP43 significantly increased (OR=8.99, 95% CI=1.07-75.72) relative to that of the GG genotype. By contrast, the combination of GA and AA genotypes was not associated with CRC risk (OR = 1.35, 95% CI = 0.92-1.99). The Del/Ins, Ins/Ins, and Del/Ins+Ins/Ins genotypes of SNP19 did not increase CRC risk relative to that of the Del/Del genotype (Del/Ins: OR = 1.33, 95% CI = 0.84-2.12; Ins/Ins: OR = 1.27, 95% CI = 0.79-2.06; Del/Ins+Ins/Ins: OR = 1.31, 95% CI = 0.84-2.04).

Table 2. Association Between SNP43, SNP19 and CRC Risk

Genotype	Cases N(%)	Controls N(%)	OR[a](95%CI)	*P*
SNP43				
GG	324(81.0)	341(85.3)	Reference	0.08
GA	69(17.2)	58(14.5)	1.25(0.83-1.86)	0.29
AA	7(1.8)	1(0.2)	8.99(1.07-75.72)	0.04
GA+AA	76(19.0)	59(14.7)	1.35(0.92-1.99)	0.13
SNP19				
Del/ Del	42(10.5)	57(14.2)	Reference	0.48
Del/Ins	188(47.0)	184(46.0)	1.33(0.84-2.12)	0.23
Ins/Ins	170(42.5)	159(39.8)	1.27(0.79-2.06)	0.32
Del/Ins+Ins/Ins	358(89.5)	343(85.8)	1.31(0.84-2.04)	0.24

[a]Adjusted by confounding factors, including family per capita annual income, family history of CRC, sitting hours per day, BMI, smoking habit, alcohol-drinking habit and tea-drinking habit

Table 3. GMDR Models of High Order Interaction on CRC Risk

Model[a]	TBA[b]	*P*	CVC[c]
Smoked meat	0.60	0.01	10/10
SNP43, Smoked meat	0.58	<0.01	6/10
SNP43, Red meat, Smoked meat	0.60	<0.01	9/10
SNP43, SNP19, Red meat, Smoked meat	0.61	<0.01	10/10

[a]Adjusted by confounding factors, including family per capita annual income, family history of CRC, sitting hours per day, BMI, smoking habit, alcohol-drinking habit and tea-drinking habit; [b]TBA, testing balance accuracy; [c]CVC, cross-validation consistency

GMDR analysis

GMDR analysis was used to analyze the high-order gene-environment interactions of SNP43, SNP19, red meat consumption, and smoked meat consumption. The best predictive models up to four orders of interaction, along with their CVCs and TBAs adjusted by the above covariates, are summarized in Table 3. The smoked meat consumption model was the best one-factor model with the highest CVC (10/10) and testing accuracy (0.60); this model was statistically significant ($P = 0.01$). Although the two-factor interaction between SNP43 and smoked meat consumption was significant ($P < 0.01$), its CVC was only 6/10. The three-factor model of SNP43, red meat consumption, and smoked meat consumption had a TBA of 0.60 ($P < 0.01$) and a CVC of 9/10. The four-factor interaction model of SNP43, SNP19, red meat

Table 4. Cumulative Effects of Risk Genotypes of SNP43, SNP19 Combined with Red Meat and Smoked Meat Consumption on CRC Risk

No.[b]	Total			Low[a]			High[a]		
	Ca/Co[c]	OR[d] (95%CI)	*P*	Ca/Co[c]	OR[d] (95%CI)	*P*	Ca/Co[c]	OR[d] (95%CI)	*P*
0	42/57	Reference	0.11	8/25	Reference	<0.01	34/32	2.99(1.16-7.71)	0.02
1	282/284	1.31(0.84-2.04)	0.24	56/117	1.39(0.58-3.31)	0.46	226/167	3.71(1.62-8.51)	<0.01
2	76/59	1.77(1.03-3.03)	0.04	18/31	1.64(0.60-4.45)	0.34	58/28	6.17(2.44-15.59)	<0.01
P_{trend}			0.04						

[a]Low red meat and low smoked meat were defined as low consumption group; otherwise, the other types were defined as high consumption group; [b]The number of risk genotypes which were defined as SNP43 GA+AA and SNP19 Del/Ins+Ins/Ins; [c]Cases and controls; d Adjusted by confounding factors, including family per capita annual income, family history of CRC, sitting hours per day, BMI, smoking habit, alcohol-drinking habit and tea-drinking habit

Table 5. Risk Estimates of CART Terminal Nodes

Node	Factors	Ca/Co[a]	OR[b] (95%CI)	P_{LR}	P_{Boot}
1	Smoked meat (Low)- Red Meat (Low)	82/173	Reference		
2	Smoked meat (Low)- Red Meat (High)- SNP43 (GA+AA)	28/6	10.50(4.01-27.54)	<0.01	<0.01
3	Smoked meat (Low)- Red Meat (High)- SNP43 (GG)- SNP19 (Del/Del)	23/19	2.16 (1.06-4.40)	0.03	0.05
4	Smoked meat (Low)- Red Meat (High)- SNP43 (GG)- SNP19 (Del/Ins+Ins/Ins)	99/105	1.75(1.17-2.60)	0.01	0.01
5	Smoked Meat (High)- SNP19 (Del/Del)- Red Meat (Low)	7/6	3.56 (1.07-11.80)	0.04	0.02
6	Smoked Meat (High)- SNP19 (Del/Del)- Red Meat (High)	4/7	1.37 (0.35-5.29)	0.65	0.61
7	Smoked Meat (High)- SNP19 (Del/Ins+Ins/Ins)- SNP43 (GG)	127/62	4.38(2.89-6.64)	<0.01	<0.01
8	Smoked Meat (High)- SNP19 (Del/Ins+Ins/Ins)- SNP43 (GA+AA) - Red Meat (Low)	10/12	2.16(0.84-5.57)	0.11	0.11
9	Smoked Meat (High)- SNP19 (Del/Ins+Ins/Ins)- SNP43 (GA+AA)- Red Meat (High)	20/10	4.56(1.94-10.75)	<0.01	<0.01

[a]Cases and controls; [b]Adjusted by confounding factors, including family per capita annual income, family history of CRC, sitting hours per day, BMI, smoking habit, alcohol-drinking habit and tea-drinking habit

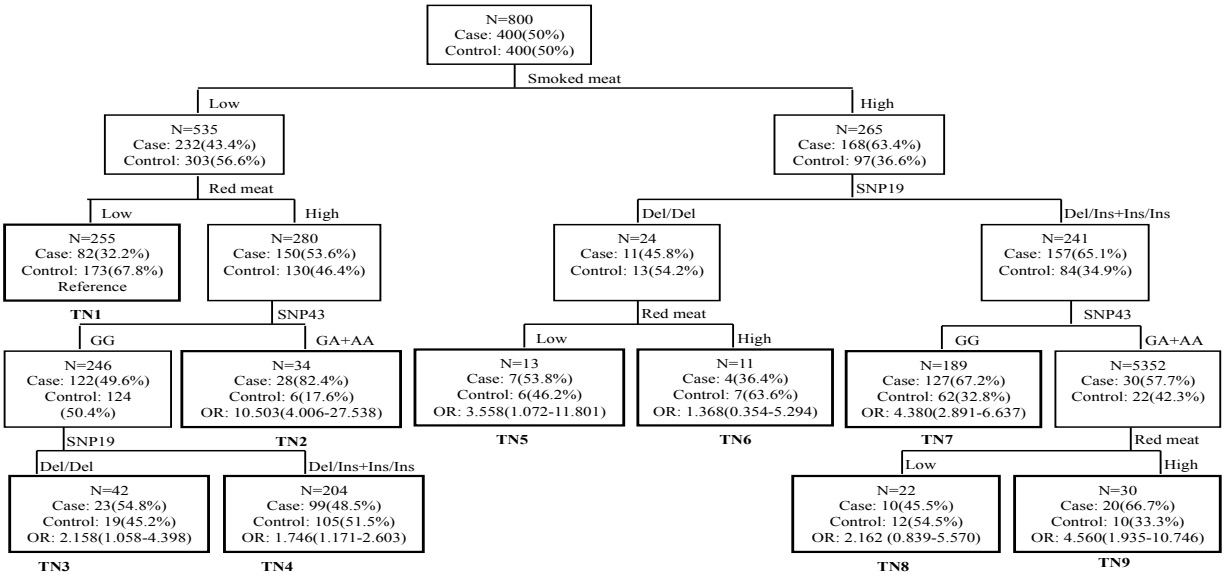

Figure 1. Classification and Regression Tree Model for SNP43, SNP19, Red Meat Consumption, and Smoked Meat Consumption. Terminal nodes are thick bordered; TN: Terminal Node

consumption, and smoked meat consumption was the best model, with a maximum CVC of 10/10 and a TBA of 0.61 ($P < 0.01$).

The cumulative effects of SNP43 and SNP19 modulated by red and smoked meat consumption were further illustrated (Table 4). Based on the results in Table 5, the genotypes of SNP43 GA+AA and SNP19 Del/Ins+Ins/Ins were considered risk genotypes. The subjects were divided into four subgroups by the number of risk genotypes. Compared with subjects that carried null risk genotypes, those with two risk genotypes had higher CRC risk (OR = 1.77, 95% CI = 1.03-3.03). A significant dosage effect with an increasing number of risk genotypes was observed with an increasing risk of CRC ($P_{trend} = 0.04$). Subjects with low red and smoked meat consumption were referred to as the low meat group, whereas those with high consumption were called the high meat group. High meat subjects with null, one, and two risk genotypes had significantly increased CRC risk (null: OR = 2.99, 95% CI = 1.16-7.71; one: OR = 3.71, 95% CI = 1.62-8.51; two: OR = 6.17, 95% CI = 2.44-15.59). In the low meat group, the number of risk genotypes was not related with CRC risk ($P > 0.05$). These results are consistent with the GMDR results, which showed that potential interactions among genes and environmental factors may affect CRC risk.

CART analysis

Figure 1 shows the CART model constructed on SNP43, SNP19, red meat consumption, and smoked meat consumption. Consistent with the best GMDR one-factor model, the first split of the root node on the decision tree was smoked meat consumption, indicating that such consumption is the strongest risk factor for CRC. The final tree contained nine terminal nodes. Terminal node 3, which consisted of the low smoked and red meat consumption groups (the least percentage of cases, 32.2%), was used as the reference to calculate the OR of the other terminal nodes(Table 5). In the high smoked meat consumption group, terminal node 9 (SNP19 Del/Ins+Ins/Ins, SNP43 GA+AA, and high red meat consumption) had the maximum risk (OR = 4.56, 95%CI = 1.94-10.75), and terminal nodes 5 (SNP19 Del/Del and low red meat consumption) and 7 (SNP19 Del/Ins+Ins/Ins and SNP43 GG) also had significant ORs (node 5: OR = 3.56, 95%CI = 1.07-11.80; node 7: OR = 4.38, 95%CI = 2.89-6.64). In the low smoked meat consumption group, the significant risk was observed for terminal nodes 2 (high red meat consumption and SNP43 GA+AA; OR = 10.50, 95% CI = 4.01-27.54), 3 (high red meat consumption, SNP43 GG, and SNP19 Del/Del; OR = 2.16, 95% CI = 1.06-4.40), and 4 (high red meat consumption, SNP43 GG, and SNP19 Del/Ins+Ins/Ins; OR = 1.75, 95%CI = 1.17-2.60). These

results indicate that potential interactions among the four factors may affect CRC risk, consistent with the GMDR results.

Discussion

CRC is a common cancer influenced by environmental factors, such as red meat consumption, smoked meat consumption, tobacco use (Zhong et al., 2013), alcohol intake (Li et al., 2011), and sedentary lifestyle (Gingras and Béliveau, 2011). By contrast, several people exposed to the same carcinogenic factors did not develop CRC. Thus, genetic background and environmental exposure interact to produce the pathological outcome of CRC. In this matched case-control study, a multi-analytic strategy combining LR, GMDR, and CART was used to systematically explore the interaction effects of SNP43, SNP19, red meat consumption, and smoked meat consumption on CRC risk. In the single-locus analysis, the GA+AA genotype of SNP43 and the Del/Ins+Ins/Ins genotype of SNP19 were marginally related to CRC risk. However, the combination of SNP43, SNP19, red meat consumption, and smoked meat consumption could significantly increase the CRC risk.

CAPN10 is traditionally considered important for IR, which is relevant to CRC risk. Except Frances et al. (2007), there is no other study has reported on the association of CAPN10 polymorphisms with CRC risk. In this study, the CAPN10 SNP43 presented a tendency to be associated with CRC risk, while the relationship of SNP19 with CRC risk was not significant. Several studies have reported the association between the genotypes of CAPN10 at SNP43 and SNP19 and IR risk. The G-to-A polymorphism of CAPN10 SNP43 is significantly associated with IR, with the strongest evidence for linkage found in the NIDDM1 region in Mexican-American sib pairs (Horikawa et al., 2000). Among Chinese, variation in SNP43 is positively associated with IR (OR=1.41, 95%CI=1.13-1.76) (Jing et al., 2012). Epidemiological studies in Japan (Shima et al., 2003) and France (Derbel et al., 2006) similarly concluded that SNP19 is involved in the pathogenesis of IR. CAPN10 is also related to several types of cancer (pancreatic, laryngeal, and gastric), and it belongs to the calpain family which contains several oncogenes and tumor suppressor genes (Liu et al., 2004; Fong et al., 2010; Moreno-Luna et al., 2011). Located at chromosome 2q 37.3, CAPN10 comprises 15 exons, spans 31 kb of genomic sequence, and encodes a 672-amino acid intracellular protease primarily expressed in the liver, skeletal muscles, and pancreatic islets (Goll et al., 2003). CAPN10 is a ubiquitous calcium-activated cysteine protease implicated in many cellular activities, including intracellular signal transduction, neuronal functions, cytoskeletal rearrangement, apoptosis, and inflammation (J Sorimachi et al., 1993). Apoptosis and inflammation have certain influences on CRC risk (Permutt et al., 2000). The over expression of human CAPN10 on transgenic mice enhances apoptosis (Johnson et al., 2004) and induces programmed cell death, which is blocked in cancer cells and thus causes abnormal tissue growth. The risk of CRC also increases with the degree of inflammation and duration of inflammatory colorectal diseases (duration/risk = 10 years/1.6%, 20 years/8.3%, and 30 years/18.4%) (Tanaka, 2012). CAPN10 is also associated with the risk of polycystic ovary syndrome (Dasgupta et al., 2012), pancreatic cancer (Fong et al., 2010), laryngeal cancer (Moreno-Luna et al., 2011), and prostate cancer (Meyer et al., 2010). The negative results of CAPN10 SNP43 and SNP19 on CRC risk in this study may due to the weak effect of single-locus or the small sample size. Further studies must be conducted in a larger sample to verify the association between CAPN10 and CRC risk.

The most significant finding of this study is the multiple gene-environment interactions consistently identified by the three statistical approaches. GMDR analysis reported the highest TBA (0.61) of combined SNP43, SNP19, red meat consumption, and smoked meat consumption for CRC risk. LR analysis indicated that CRC risk was associated with SNP43 and SNP19 in a gene dose-dependent manner. High meat subjects with two risk genotypes significantly increased CRC risk (OR = 6.17, 95% CI = 2.44-15.59), relative to low meat subjects with null risk genotype. CART analysis found a consistent interaction result among the four factors, with an OR of 4.56 (95%CI = 1.94-10.75). Combined with high consumption of red and smoked meat, the cumulative effect of SNP43 and SNP19 on CRC risk significantly strengthened, suggesting that the four-factor model is the best model for assessing CRC risk in our study. These results are biologically plausible because red meat and smoked meat are related to IR and CRC. Red meat is rich in iron, which directly and causally affects T2DM pathogenesis mediated by both β cell failure and IR (Simcox and McClain, 2013). A high-red meat diet raises fat intake and thus increases the plasma insulin concentrations (Chaolu et al., 2011). An experiment on mice concluded that a high-fat diet increases insulin levels and associated enzyme expression in skeletal muscles (Barnea et al., 2006). High red meat consumption is a significant CRC risk factor (WCRF, 2007; Hu et al., 2013), as evidenced by the promotive effect of high red meat consumption on CRC risk in this study. One mechanism for the link between high-red meat diets and CRC is attributed to high fat levels (Tang et al., 2012). Animal fats, particularly from red meat, are a risk factor for colon cancer among women (Macrae, 1993). Epidemiological and experimental studies support the hypothesis that heme iron in red meat causes CRC. This effect may be attributed to the catalytic effect of heme iron on the endogenous formation of carcinogenic N-nitroso compounds and cytotoxic and genotoxic aldehydes by lipoperoxidation (Bastide et al., 2011). Smoked meat significantly influences the promotion of colon carcinogenesis (Santarelli et al., 2010). During the preparation of smoked meat, N-nitroso compounds are formed in the meat as it is pickled with salt, and heterocyclic amines are also formed as the meat is cooked at high temperature; both compounds are strong carcinogens for CRC (Bingham, 1999; Knize et al., 1999; Tang et al., 2012). SNP43 is associated with intra-abdominal fat, a large waist circumference, and increased serum cholesterol, all of which can be caused by high consumption of red and smoked meat. This association

confirms the link between red and smoked meat and SNP43 signaling. CAPN10 SNP19 may significantly influence the regulation of thermogenesis by reducing beta(3)-adrenoceptor function. Thermogenesis is closely related to the consumption of red and smoked meat. Therefore, CAPN10 SNP43 and SNP19 may affect the risk of CRC together with high consumption of red and smoked meat.

This study has several limitations. First, the lifestyle and dietary habits of the subjects were collected in reference to 10 years before the disease diagnosis. Thus, recall bias was difficult to control completely. We trained interviewers and recruited newly diagnosed CRC patients to reduce the bias. Second, the number of SNPs related to CAPN10 was limited, whereas there were many haplotypes proven to be associated with IR. Third, the sample size of our study was relatively small to detect a significant association between SNPs of CAPN10 and CRC risk. The statistical powers for SNP43 and SNP19 in our study were less than 80%. In our next study, we will explore more SNPs of CAPN10 in a larger sample size to assess the effect of CAPN10 on CRC risk and determine the susceptibility of CAPN10 SNPs to CRC.

In conclusion, the single locus of CAPN10 SNP43 and SNP19 is not enough to significantly increase the CRC risk. However, SNP43, SNP19, red meat consumption, and smoked meat consumption potentially work together to influence CRC risk. It implied that subjects carried CAPN10 SNP43 GA+AA and SNP19 Del/Ins+Ins/Ins should intake lower red and smoked meat than those carried null risk genotypes.

Acknowledgements

The authors wish to thank the participants for their efforts in completing the study. The study was supported by the Doctoral Fund of Ministry of Education of China (Grant Number 20090181120019) and the National Natural Science Foundation of China (Grant Number 81102196).

References

Barnea M, Shamay A, Stark AH, Madar Z (2006). A high-fat diet has a tissue-specific effect on adiponectin and related enzyme expression. *Obesity (Silver Spring)*, **14**, 2145-53.

Basen-Engquist K, Chang M (2011). Obesity and cancer risk: recent review and evidence. *Curr Oncol Rep*, **13**, 71-6.

Bastide NM, Pierre FH, Corpet DE (2011). Heme iron from meat and risk of colorectal cancer: a meta-analysis and a review of the mechanisms involved. *Cancer Prev Res (Phila)*, **4**, 177-84.

Bingham SA (1999). High-meat diets and cancer risk. *Proc Nutr Soc*, **58**, 243-8.

Chaolu H, Asakawa A, Ushikai M, et al (2011). Effect of exercise and high-fat diet on plasma adiponectin and nesfatin levels in mice. *Exp Ther Med*, **2**, 369-73.

Cowey SL, Quast M, Belalcazar LM, et al (2005). Abdominal obesity, insulin resistance, and colon carcinogenesis are increased in mutant mice lacking gastrin gene expression. *Cancer*, **103**, 2643-53.

Dasgupta S, Sirisha PV, Neelaveni K, et al (2012). Association of CAPN10 SNPs and haplotypes with polycystic ovary syndrome among South Indian Women. *PLoS One*, **7**, e32192.

de Oliveira Otto MC, Alonso A, Lee DH, et al (2012). Dietary intakes of zinc and heme iron from red meat, but not from other sources, are associated with greater risk of metabolic syndrome and cardiovascular disease. *J Nutr*, **142**, 526-33.

Derbel S, Doumaguet C, Hubert D, et al (2006). Calpain 10 and development of diabetes mellitus in cystic fibrosis. *J Cyst Fibros*, **5**, 47-51.

Djiogue S, Nwabo Kamdje AH, et al (2013). Insulin resistance and cancer: the role of insulin and IGFs. *Endocr Relat Cancer*, **20**, R1-R17.

Ferlay J, Shin HR, Bray F, et al (2010). GLOBOCAN 2008 v1.2, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. Lyon, France: International Agency for Research on Cancer; Available from: http://globocan.iarc.fr, accessed on 8/9/2012.

Fong PY, Fesinmeyer MD, White E, et al (2010). Association of diabetes susceptibility gene calpain-10 with pancreatic cancer among smokers. *J Gastrointest Cancer*, **41**, 203-8.

Frances CP, Conde MC, Saez ME, et al (2007). Identification of a protective haplogenotype within CAPN10 gene influencing colorectal cancer susceptibility. *J Gastroenterol Hepatol*, **22**, 2298-302.

Gingras D, Béliveau R (2011). Colorectal cancer prevention through dietary and lifestyle modifications. *Cancer Microenviron*, **4**, 133-9.

Giovannucci E (1995). Insulin and colon cancer. *Cancer Causes Control*, **6**, 164-79.

Goll DE, Thompson VF, Li H, et al (2003). The calpain system. *Physiol Rev*, **83**, 731-801.

Hanis CL, Boerwinkle E, Chakraborty R, et al (1996). A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat Genet*, **13**, 161-6.

Horikawa Y, Oda N, Cox NJ, et al (2000). Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet*, **26**, 163-75.

Hu X, Yuan P, Yan J, et al (2013). Gene Polymorphisms of ADIPOQ +45T>G, UCP2 -866G>A, and FABP2 Ala54Thr on the Risk of Colorectal Cancer: A Matched Case-Control Study. *PLoS One*, **8**, e67275.

J Sorimachi H, Ishiura S, Suzuki K (1993). A novel tissue-specific calpain species expressed predominantly in the stomachcomprises two alternative splicing products with and without Ca(2+)-binding domain. *Biol Chem*, **268**, 19476-82.

Jing C, Xueyao H, Linong J (2012). Meta-analysis of association studies between five candidate genes and type 2 diabetes in Chinese Han population. *Endocrine*, **42**, 307-20.

Johnson JD, Han Z, Otani K, et al (2004). RyR2 and calpain-10 delineate a novel apoptosis pathway in pancreatic islets. *J Biol Chem*, **279**, 24794-802.

Knize MG, Salmon CP, Pais P, Felton JS (1999). Food heating and the formation of heterocyclic aromatic amine and polycyclic aromatichydrocarbon mutagens/carcinogens. *Adv Exp Med Biol*, **459**, 179-93.

Komninou D, Ayonote A, Richie JP Jr, Rigas B (2003). Insulin resistance and its contribution to colon carcinogenesis. *Exp Biol Med (Maywood)*, **228**, 396-405.

Kotnis A, Sarin R, Mulherkar R (2005). Genotype, phenotype and cancer: Role of low penetrance genes and environment in tumour susceptibility. *J Biosci*, **30**, 93-102.

Kuhajda FP (2008). AMP-activated protein kinase and human cancer: cancer metabolism revisited. *Int J Obes (Lond)*, **32**, S36-41.

Lajous M, Tondeur L, Fagherazzi G, et al (2012). Processed and unprocessed red meat consumption and incident type 2 diabetes among French women. *Diabetes Care*, **35**, 128-30.

Leng SL, Leeding KS, Gibson PR, Bach LA (2001). Insulin-like growth factor-II renders LIM 2405 human colon cancer cells resistant to butyrate-induced apoptosis: a potential mechanism for colon cancer cell survival in vivo. Carcinogenesis, **22**, 1625-31.

Li Y, Yang H, Cao J (2011). Association between alcohol consumption and cancers in the Chinese population--a systematic review and meta-analysis. *PLoS One*, **6**, e18776.

Liu X, Van Vleet T, Schnellmann RG (2004). The role of calpain in oncotic cell death. *Annu Rev Pharmacol Toxicol*, **44**, 349-70.

Lou XY, Chen GB, Yan L, et al (2007). A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am J Hum Genet,* **80**, 1125-37.

Macrae FA (1993). Fat and calories in colon and breast cancer: from animal studies to controlled clinical trials. *Prev Med*, **22**, 750-66.

McKeown-Eyssen G (1994). Epidemiology of colorectal cancer revisited: are serum triglycerides and/or plasma glucose associated with risk? *Cancer Epidemiol Biomarkers Prev*, **3**, 687-95.

Meyer TE, Boerwinkle E, Morrison AC, et al (2010). Diabetes genes and prostate cancer in the Atherosclerosis Risk in Communities study. *Cancer Epidemiol Biomarkers Prev*, **19**, 558-65.

Moreno-Luna R, Abrante A, Esteban F, et al (2011). Calpain 10 gene and laryngeal cancer: a survival analysis. *Head Neck*, **33**, 72-6.

Noto H, Tsujimoto T, Sasazuki T, Noda M (2011). Significantly increased risk of cancer in patients with diabetes mellitus: a systematic review and meta-analysis. *Endocr Pract*, **17**, 616-28.

Ortiz AP, Thompson CL, Chak A, et al (2012). Insulin resistance, central obesity, and risk of colorectal adenomas. *Cancer*, **118**, 1774-81.

Permutt MA, Bernal-Mizrachi E, Inoue H (2000). Calpain 10: the first positional cloning of a gene for type 2 diabetes? *J Clin Invest*, **106**, 819-21.

Pollak M (2000). Insulin and insulin-like growth factor signalling in neoplasia. *Nature Reviews Cancer*, **8**, 915-28.

Sáez ME, González-Sánchez JL, Ramírez-Lorca R, et al (2008). The CAPN10 gene is associated with insulin resistance phenotypes in the Spanishpopulation. *PLoS One*, **3**, e2953.

Santarelli RL, Naud N, Taché S, et al (2013). Calcium inhibits promotion by hot dog of 1,2-dimethylhydrazine-induced mucin-depleted foci in rat colon. *Int J Cancer*, **133**, 2533-41.

Santarelli RL, Vendeuvre JL, Naud N, et al (2010). Meat processing and colon carcinogenesis: cooked, nitrite-treated, and oxidized high-heme smoked meat promotes mucin-depleted foci in rats. *Cancer Prev Res (Phila)*, **3**, 852-64.

Shima Y, Nakanishi K, Odawara M, et al (2003). Association of the SNP-19 genotype 22 in the calpain-10 gene with elevated body mass index and hemoglobin A1c levels in Japanese. *Clin Chim Acta*, **336**, 89-96.

Simcox JA, McClain DA (2013). Iron and diabetes risk. *Cell Metab*, **17**, 329-41.

Tanaka T (2012). Development of an inflammation-associated colorectal cancer model and its application for research on carcinogenesis and chemoprevention. *Int J Inflam*, **2012**, 658786.

Tang FY, Pai MH, Chiang EP (2012). Consumption of high-fat diet induces tumor progression and epithelial-mesenchymal transition of colorectal cancer in a mouse xenograft model. *J Nutr Biochem*, **23**, 1302-13.

Tuncman G, Hirosumi J, Solinas G, et al. (2006). Functional in vivo interactions between JNK1 and JNK2 isoforms in obesity and insulin resistance. *Proc Natl Acad Sci USA*, **103**, 10741-6.

Turner MD, Cassell PG, Hitman GA (2005). Calpain-10: from genome search to function. *Diabetes Metab Res Rev*, **21**, 505-14.

Wang W, Guan KL (2009). AMP-activated protein kinase and cancer. *Acta Physiol (Oxf)*, **196**, 55-63.

WCRF (2007). Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. World Cancer Research Fund & American Institute for Cancer Research, Washington, DC.

Wilkin TJ (2009). The accelerator hypothesis: a review of the evidence for insulin resistance as the basis for type I as well as type II diabetes. *Int J Obes (Lond)*, **33**, 716-26.

Zeyda M, Stulnig TM (2009). Obesity, inflammation, and insulin resistance-a mini-review. *Gerontology*, **55**, 379-86.

Zhao P, Zhang WQ (2011). Chinese tumor registration report in 2010. Beijing, China: Military medical science press.

Zhong R, Liu L, Zou L, et al (2013). Genetic variations in the TGFβ signaling pathway, smoking and risk of colorectal cancer in a Chinese population. *Carcinogenesis*, **34**, 936-42.