UNIVERSIDADE D COIMBRA FACULDADE DE MEDICINA

MESTRADO INTEGRADO EM MEDICINA – TRABALHO FINAL

RUI FILIPE DA SILVA GOMES

A single nucleotide polymorphism in IREB2 gene is associated with colorectal cancer: a case-control study

ARTIGO CIENTÍFICO

ÁREA CIENTÍFICA DE ONCOLOGIA

Trabalho realizado sob a orientação de:

PROFESSOR DOUTOR JOSÉ MANUEL BORGES NASCIMENTO COSTA

PROFESSORA DOUTORA ANA BELA SARMENTO ANTUNES CRUZ RIBEIRO

ABRIL/2019

A single nucleotide polymorphism in IREB2 gene is associated with colorectal cancer: a case-control study

Rui Filipe da Silva Gomes[a,b], Joana Margarida Verdasca Jorge[b,c,d], Raquel Fernanda da Silva Alves[b,c,d], Ana Cristina Pereira Gonçalves[b,c,d], Ana Bela Sarmento Antunes Cruz Ribeiro[b,c,d,e], José Manuel Borges Nascimento Costa[c,d,f]

a. Medical Student, Faculty of Medicine, University of Coimbra, Portugal
b. Laboratory of Oncobiology and Hematology (LOH) and University Clinic of Hematology/Faculty of Medicine, University of Coimbra, Portugal
c. Coimbra Institute for Clinical and Biomedical Research (iCBR) – Group of Environment, Genetics and Oncobiology (CIMAGO) - Faculty of Medicine, University of Coimbra, Portugal
d. Center for Innovative Biomedicine and Biotechnology (CIBB), Coimbra, Portugal
e. Clinical Hematology Department/Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
f. University Clinic of Oncology/Faculty of Medicine, University of Coimbra, Portugal

Correspondence:

Professor Doutor José Manuel Borges Nascimento Costa
Subunidade 1, Azinhaga de Santa Comba, Celas, Portugal
3000-548 – Coimbra
E-mail address: uc12934@uc.pt

TABLE OF CONTENTS

TABLE INDEX

FIGURE INDEX

ABSTRACT

Colorectal cancer (CRC) is one of the most common and lethal cancers worldwide. CRC risk is determined by a complex interaction between environmental exposures and genetic variants. A growing body of evidence has proved that iron overload and dietary iron contribute to CRC development. Considering the fundamental role of iron-regulatory protein 2 (IRP2) in the regulation of intracellular iron homeostasis, it is reasonable that some polymorphisms in *IREB2* gene, which encode IRP2, could be involved in CRC carcinogenesis. In this context, we investigated the role of *IREB2* variant rs17483548 in CRC susceptibility, location, staging, and prognosis, aiming to identify new potential risk factors and/or prognostic markers.

A hospital-based case-control study was conducted with 83 CRC patients and 176 healthy controls. The DNA from patients and controls was extracted from whole blood samples and the *IREB2* variant rs17483548 corresponding genomic region was amplified by tetra-primer ARMS-PCR assay. Allele and genotype frequencies were determined and compared between CRC patients and controls. The Fisher's exact test was used to compare allele frequencies, while the logistic regression analysis was performed to evaluate the possible associations between *IREB2* variant rs17483548 genotypes and CRC development, location and stage. The overall survival was analyzed by Kaplan-Meier method.

The results showed that *IREB2* GA (OR = 0.522, 95% CI 0.287 - 0.951, $p = 0.034$) and AA (OR = 0.361, 95% CI 0.167 - 0.782, $p = 0.010$) genotypes are associated with a decreased susceptibility for CRC, while GG genotype showed a 2-fold increase in the CRC risk (OR = 2.133, 95% CI 1.212 - 3.755, $p = 0.009$). Moreover, GA and AA genotypes were associated with an increased predisposition for locoregional CRC which can indicate that these genotypes are associated with low aggressive tumors. In contrast, GG genotype patients had a lower predisposition for locoregional CRC but did not show any statistical significant association with more advanced tumor stages. The association analysis between *IREB2* genotypes and CRC location also showed a decreased predisposition for right and transverse colon neoplasms among GG genotype patients. The overall survival was not influenced by *IREB2* polymorphism (HR = 1.269; 95% CI 0.615 - 2.620; $p = 0.519$).

Altogether, these results suggest that *IREB2* variant rs17483548 may play an important role in CRC susceptibility and staging. However, more studies are needed to characterize better the impact of this single nucleotide polymorphism in CRC.

KEYWORDS

Colorectal neoplasms; IREB2; single nucleotide polymorphism; iron-regulatory proteins; iron metabolism.

RESUMO

O cancro colorretal (CCR) é um dos cancros mais comuns e mortais em todo o mundo. O seu risco é determinado por uma complexa interação entre fatores ambientais e variantes genéticas. Um número crescente de estudos tem demonstrado que o ferro da dieta e a sobrecarga de ferro contribuem para o desenvolvimento de CCR. Tendo em conta o papel fundamental da proteína IRP2 (*iron-regulatory protein 2*) na regulação da homeostasia do ferro intracelular, é expectável que os polimorfismos do gene *IREB2*, que codifica a IRP2, possam estar envolvidos na carcinogénese do CCR. Neste contexto, decidimos investigar o papel da variante genética rs17483548 do gene *IREB2* na susceptibilidade, localização, estadiamento e prognóstico do CCR, com o objetivo de identificar potenciais fatores de risco e biomarcadores de prognóstico.

Realizou-se um estudo de caso-controlo de base hospitalar com 83 doentes e 176 indivíduos saudáveis. O DNA dos doentes e dos controlos foi extraído a partir de amostras de sangue total e a região genómica correspondente à variante genética rs17483548 do gene IREB2 foi amplificada pela técnica de *tetra-primer ARMS-PCR*. As frequências alélicas e genotípicas foram determinadas e comparadas entre os doentes com CCR e os controlos. O teste exato de Fisher foi utilizado para comparar as frequências alélias, enquanto a análise de regressão logística foi usada para avaliar as possíveis associações entre os genótipos da variante genética rs17483548 do gene IREB2 e o desenvolvimento, a localização e o estadio do CCR. A sobrevivência global foi analisada pelo método de Kaplan-Meier.

Os resultados mostraram que os genótipos GA (OR = 0.522, 95% CI 0.287 - 0.951, p = 0.034) e AA (OR = 0.361, 95% CI 0.167 - 0.782, p = 0.010) do gene *IREB2* estão associados a menor suscetibilidade para o CCR, enquanto o genótipo GG (OR = 2.133, 95% CI 1.212 - 3.755, p = 0.009) mostrou aumentar em 2 vezes o risco deste cancro. Além disso, os genótipos GA e AA foram associados a um aumento da predisposição para o CCR de estadio locoregional, o que poderá indicar que estes genótipos se associam a tumores menos agressivos. Em contraste, os doentes com genótipo GG apresentaram uma menor predisposição para o CCR de estadio locoregional mas não mostraram nenhuma associação estatisticamente significativa com os tumores de estadio mais avançado. A análise de associação entre os genótipos do *IREB2* e a localização do CCR também revelou uma menor predisposição para as neoplasias do cólon direito e transverso entre os doentes com

genótipo GG. A sobrevivência global não foi influenciada pelo polimorfismo do gene *IREB2* (HR = 1.269; 95% CI 0.615 - 2.620; $p = 0.519$).

Em conjunto, estes resultados sugerem que a variante genética rs17483548 do gene *IREB2* poderá desempenhar um papel importante na susceptibilidade e no estadiamento do CCR. Porém, mais estudos serão necessários para caracterizar melhor o impacto deste polimorfismo de nucleótido único no CCR.

PALAVRAS-CHAVE

Cancro colorretal; IREB2; polimorfismo de nucleótido único; proteínas reguladoras do ferro; metabolismo do ferro.

ABBREVIATIONS

AJCC, The American Joint Committee on Cancer

ARMS-PCR, Amplification-refractory mutation system-polymerase chain reaction

CI, Confidence interval

CRC, Colorectal cancer

DNA, Deoxyribonucleic acid

dNTPs, Deoxynucleoside triphosphate

EDTA, Ethylenediaminetetraacetic acid

EPE, Entidade pública empresarial

FMUC, Faculty of Medicine, University of Coimbra, Portugal

GWAS, Genome-wide association study

GRS, Genetic risk score

HDFF, Hospital Distrital da Figueira da Foz

HDI, Human development index

HWE, Hardy-Weinberg equilibrium

ICD-10, 10th revision of the International Statistical Classification of Diseases and Related Health Problems

IRE, Iron responsive element

IREB2, Iron-responsive element binding-protein 2

IRP, Iron-regulatory protein

IRP1, Iron-regulatory protein 1

IRP2, Iron-regulatory protein 2

LOH, Laboratory of Oncobiology and Hematology

MAF, Minor allele frequency

mRNA, Messenger ribonucleic acid

OR, Odds ratio

OS, Overall survival

PCR, Polymerase chain reaction

RECIST, Response evaluation criteria in solid tumors

ROS, Reactive oxygen species

SNP, Single nucleotide polymorphism

TfR, Transferrin receptor

TNM, TNM classification of malignant tumors

UTR, Untranslated region

USA, United States of America

INTRODUCTION

Colorectal cancer (CRC) is the third most common and the second most lethal cancer worldwide. In 2018, this type of cancer accounted for about 1.8 million new cases and 881 000 deaths. The majority of all cases occurred in more developed regions and the highest mortality rates were observed in Central and Eastern Europe.[1,2] However, CRC incidence and mortality are increasing rapidly in many countries with medium-to-high human development index (HDI). At the same time, some highest indexed HDI countries are achieving stabilization or decline in the incidence and mortality rates, probably due to better practices in CRC prevention, screening, and treatment.[3]

CRC risk is determined by a complex interaction between environmental exposures and genetic variants.[4] Several risk factors have been well-established, including older age, male sex, family history of CRC, inflammatory bowel disease, and some common elements of the western lifestyle (high consumption of red and processed meat, excessive alcohol consumption, smoking, physical inactivity, diabetes, and obesity). On the other hand, some preventive factors include physical activity, regular use of aspirin, hormone replacement therapy, and endoscopy examination with the removal of precancerous lesions.[5,6]

In addition, numerous population studies demonstrated a positive association between CRC and dietary iron or elevated iron levels.[7–11] In fact, iron could have a role in carcinogenesis due to its ability to lose and gain electrons. This attribute enables iron to participate in free radical-generating reactions, such as Fenton reaction, in which ferrous iron reacts with hydrogen peroxide to produce a reactive oxygen species (ROS) termed hydroxyl radical.[12] Free radicals react with proteins, lipids, and DNA causing damage in these biological molecules. In living organisms, DNA damage is repaired by a range of mechanisms. But, when DNA damage repair does not occur correctly, the damage accumulation leads to genetic instability which can induce carcinogenesis.[13] Moreover, iron is a nutrient that promotes tumor cell proliferation and growth.[12]

Recent insights demonstrated that cancer cells are more dependent on iron for growth and more susceptible to iron depletion than normal cells, a phenomenon called *iron addiction*. Other studies showed an increased intracellular iron pool in some cancer cells caused by dysregulation in iron uptake, storage, and efflux.[14] These alterations are consistent with another emerging concept that tumors create an iron-rich microenvironment to circumvent the limited systemic iron availability and to modulate the immune response.[12,15]

Iron-regulatory proteins (IRP1 and IRP2) regulate iron homeostasis and use. IRP1 and IRP2 are cytosolic RNA-binding proteins that bind to iron-responsive elements (IREs), stem-loop structures located in either the 5' or 3' untranslated regions (UTRs) of specific mRNAs.[16] IRP2 is encoded by *IREB2* gene and differs from IRP1 because its activity is only determined by its concentration, whereas IRP1 activity is primarily regulated by an unusual iron-sulfur cluster switch.[17,18] When iron levels are high, IRP2 undergoes proteasomal degradation. Other physiological stimuli are also involved in the regulation of the IRP2 concentration, including ROS, nitric oxide, and hypoxia.[19] The available IRP2 binds IREs located in the 5'-UTR of the mRNAs encoding ferritin and erythroid aminolevulinic acid synthase, inhibiting the translation of these mRNAs. Simultaneously, IRP2 binds IREs located in the 3'-UTR of the transferrin receptor (TfR) mRNA, stabilizing it.[16] Therefore, under iron excess conditions, it is expected that IRP2 goes under degradation, resulting in TfR suppression and ferritin increased expression. However, iron-rich tumors, including CRC, do not express such profile.[20] It suggests a possible involvement of IRP2 in tumorigenesis, an idea supported by several studies.[20–26]

Complex diseases, such as CRC, arise from the combination of exposures to environmental factors with acquired mutations and germline variants, including single nucleotide polymorphisms (SNP) in more than one gene. This factors modulate disease susceptibility, severity, clinical manifestations, and treatment responses.[27] The present study aimed to investigate the influence of a SNP in *IREB2* gene (rs17483548) on the CRC susceptibility, location, staging, and prognosis, in order to identify new potential risk factors and/or prognostic markers. The identification of genetic variants associated with CRC may be of great relevance to understand better its molecular mechanisms and to develop new screening or preventive strategies and therapeutic approaches.

MATERIALS AND METHODS

Ethical Statement

The present study was conducted in accordance with the Helsinki declaration of 1975 (revised in 2004 and 2008) and the protocols were approved by the Ethics Committee of FMUC - Faculty of Medicine, University of Coimbra, Portugal. All participants were informed about the study and signed an informed consent form before study enrollment.

Study Design and Population

We performed a hospital-based case-control study in blood samples of 83 patients with colorectal cancer and 176 healthy controls. All cases and controls were enrolled from one hospital of the central region of Portugal – "Hospital Distrital da Figueira da Foz, EPE (HDFF, EPE)" – from August 2009 to January 2013. The diagnosis was based on the biopsy of suspicious tumoral lesions and subsequent identification of CRC through anatomopathological examination. Afterward, the TNM classification and tumor staging were assessed for all diagnosed patients according to the 7th edition of The American Joint Committee on Cancer (AJCC) guidelines, as revised in 2010. All tumors were also classified according to ICD-10 criteria (version for 2010). The overall survival (OS) was selected as study endpoint and was measured from the date of diagnosis. The OS endpoints used in the present study were deceased or alive at the moment of the last contact with the medical team (patients who were still alive were censored). Controls were selected among healthy blood donors with no personal history of cancer or inflammatory bowel disease, in the same hospital**.** In order to control the effect of confounders, matching based on gender and age was carried out between cases and controls. The demographic and clinical data about all patients and controls were collected during a personal interview and through consultation of the medical records, after study enrollment.

Gene and SNP Selection

IREB2 gene was selected based on its fundamental role on iron homeostasis regulation and due to its documented relation with some solid neoplasms, such as breast, prostate, and lung cancers.[23–26] The selection of *IREB2* variant rs17483548 was based on the following criteria: (1) minor allele frequency (MAF) ≥10% in Caucasians according to published

literature or public databases; (2) SNP was previously validated; (3) promising relevance to cancer development. Pubmed (accessible in https://www.ncbi.nlm.nih.gov/pubmed/) and dbSNP (accessible in https://www.ncbi.nlm.nih.gov/SNP/) databases were consulted to obtain the necessary information. Table 1 shows all important details about this *IREB2* genetic variant.

Table I. Most relevant information about *IREB2* variant rs17483548.

Gene symbol*	***IREB2***
dbSNP	rs17483548
Chromosomal position	15:78437971
Alleles	G>A
Variant	g.49520870G>A
mRNA position	Intron
Global MAF¥ (allele)	0.2220 (A)
IBS MAF§ (allele)	0.4393 (A)

dbSNP, single nucleotide polymorphism database; mRNA, messenger ribonucleic acid; MAF, minor allele frequency; IBS, Iberian Population in Spain; *IREB2*, iron-responsive element binding-protein 2; SNP, single nucletide polymorphism. The data exposed in this table were collected from the following databases: *HUGO Gene Nomenclature Committee (HGNC); ¥TOPMed (Trans-Omics for Precision Medicine); §1000Genomes.

DNA Extraction and SNP Genotyping

Genomic DNA was extracted from the whole blood samples, using the salting-out method as previously described by Miller *et al.*[28] The extracted DNA was diluted in nuclease-free water and stored at -80 ºC before quantification using a NanoDrop ND-100 spectrophotometer (NanoDrop Technologies, Wilmington, DE). Finally, 100 ng of each DNA sample was used in the genotyping assay.

The *IREB2* variant rs17483548 was genotyped using tetra-primer ARMS-PCR assay**.** The primers used in this ARMS-PCR were designed with BatchPrimer3 1.0 software (available in http://probes.pw.usda.gov/batchprimer3/) with the following sequences: forward inner primer (G allele) – 5'-CGA ATT CAT GAC CCC AAT AGA AAA CTT CAG-3', reverse inner primer (A allele) – 5'-CAG CGA TCC GTT ACT TAG TTG CGT CT-3', forward outer primer – 5'-TCT CCT TGA GCT CTT TCT CCT AGC AGT TC-3', and reverse outer primer – 5'-AGA TCG

TCG GAC AGG AAA ACA AAG AA-3'. Amplifications were performed in a final volume of 20 µL containing 100 ng of template DNA, 1.5X of PCR buffer, 0.8 mM of dNTPs, 2 mM of $MgCl_2$, 0.16 µM of reverse inner and forward outer primers, 0.08 µM of forward inner and reverse outer primers, and 2 U of Taq DNA polymerase enzyme. PCR amplification conditions were programmed on a thermocycler (T100™ Thermal Cycler, Bio-Rad, USA) as follows: 5 min of denaturation step at 95 ºC followed by 35 amplification cycles at 95 ºC for 30 sec, 61 ºC for 45 sec and 72 ºC for 45 sec, and finishing with 10 min of a final extension step at 72 ºC.

The PCR products were analyzed by electrophoresis on a 3% agarose gel stained with Green Safe dye, using three samples containing the three possible genotypes as positive controls and distilled water as no template control. The G allele has 203 base pairs of size, while the A allele has 135 bp. The product of the two outer primers has 282 bp. All genotypes were scored independently by two different reviewers. Figure 1 shows an electrophoresis strip containing the three possible genotypes (GG, GA, and AA).

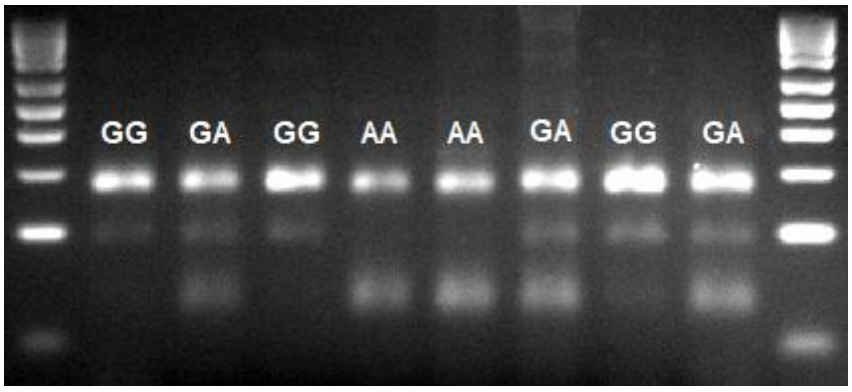

Figure 1. Representative example of an electrophoresis strip containing the three possible genotypes (GG, GA, and AA) of the *IREB2* variant rs17483548. This electrophoresis was performed on a 3% agarose gel with the amplification products of the tetra-primer ARMS-PCR. The fragments with 282 base pairs of size correspond to the products of the two outer primers. The GG genotype has one fragment with 203 bp, the AA genotype has one fragment with 135 bp and the GA genotype has two fragments with 203 bp and 135 bp each.

Statistical Analysis

Statistical analyses were performed using IBM SPSS version 25 (IBM, Armonk, NY, USA) and GraphPad Prism version 6 (GraphPad Software, San Diego, CA, USA) with the help of Microsoft Office Excel 2010 (Microsoft Corporation, Redmond, WA, USA). Normality and

differences in demographic characteristics between cases and controls were tested using the nonparametric Mann-Whitney U test for age and the chi-squared test for gender. Genotype and allele frequencies were calculated by direct counting. The agreement with Hardy-Weinberg equilibrium (HWE) was assessed in both studied groups using Arlequin software version 3.5.1.2 (available in http://cmpg.unibe.ch/software/arlequin3512/).

To explore the possible association between *IREB2* variant rs17483548 and CRC, the minor allele was compared with the major allele as reference, using the Fisher's exact test. Furthermore, we evaluated the possible associations between *IREB2* genotypes and CRC development, location and stage, using logistic regression analysis, by calculating the odds ratio (OR) and its 95% confidence interval (CI). The homozygous genotype for the major allele was considered the reference. This analysis accounted to four genetic models: (1) codominant model (each genotype was compared with homozygous controls for the major allele); (2) dominant model (minor allele carriers were compared with homozygous controls for the major allele); (3) over-dominant model (heterozygous were compared with homozygous controls for the major and the minor alleles); (4) recessive model (homozygous for the minor allele were compared with controls carrying the major allele).

In order to achieve adequate statistical power and to simplify the statistical analysis, patients were divided into the following three groups according to CRC location: (1) neoplasms of the right and transverse colon (which include all tumors located in the caecum, appendix, ascending colon, hepatic flexure, and transverse colon); (2) neoplasms of the left colon (which include all tumors of the splenic flexure, descending colon, sigmoid colon, and rectosigmoid junction); (3) neoplasms of the rectum. For the same reasons, all patients were divided into three groups considering the obtained tumor stage according to the 7th edition of the AJCC guidelines: (1) locoregional (which corresponds to stages I and IIA-IIC); (2) lymph nodes involvement (which corresponds to stages IIIA-IIIC); (3) distant metastasis (which corresponds to stages IVA-IVB).

The overall survival (OS) of patients, stratified according to their genotypes, was analyzed by Kaplan-Meier method. Log-rank test was used to assess differences in survival. The hazard ratio (HR) and its 95% CI were estimated by a multivariate Cox proportional hazards model.

All statistical analyses performed in the present study were two-sided and a p-value <0.05 was considered significant.

RESULTS

Characterization of CRC Patients and Controls

The present study enrolled 83 patients with CRC diagnosed after anatomopathological examination, of whom 55 (66.3%) were males and 28 (33.7%) were females. The patients' median age was 71 years old, ranging from 28 years to 91 years. The control group was composed of 176 healthy individuals, including 115 males (65.3%) and 61 females (34.7%) with a median age of 73 years old (ranging from 21 years to 95 years). In order to confirm adequate matching between CRC patients and controls, we assessed differences in the baseline demographic characteristics between these two groups. There were no statistically significant differences in age (p=0.870) and gender (p=1.000), which indicates an adequate group matching. Table 2 summarizes the basic demographic characteristics of both patients and controls groups.

According to the ICD-10 criteria (version for 2010), the diagnosed tumors were located in the following sites: caecum (n = 9; 10.8%), ascending colon (n = 8; 9.6%), hepatic flexure (n = 3; 3.6%), transverse colon (n = 3; 3.6%), splenic flexure (n = 1; 1.2%), descending colon (n = 2; 2.4%), sigmoid colon (n = 18; 21.7%), rectosigmoid junction (n = 5; 6.0%), and rectum (n = 30; 36.1%). Four patients (4.8%) had tumors diagnosed as CRC but not classified according to the ICD-10 criteria. However, two of those patients had documented tumor lesions in the right colon and the other two patients had tumors in the left colon, as registered in the medical records. Considering the previous distribution and the information collected about those four participants, we organized all patients into three groups according to CCR location: (1) neoplasms of the right and transverse colon (n = 25; 30.5%); (2) neoplasms of the left colon (n = 28; 34.1%); (3) neoplasms of the rectum (n = 30; 36.6%).

In addition, all patients were stratified according to the 7th edition of the AJCC guidelines. The tumor staging allowed us to distribute the patients into four stage groups: stage I (n = 18; 21.7%), stage II (n = 24; 28.9%), stage III (n = 22; 26.5%), and stage IV (n = 16; 19.3%). The remaining three patients (3.6%) had an unknown tumor stage. Considering this information, we organized all patients with a known tumor stage into the three groups previously described to simplify the statistical analysis: (1) locoregional (n = 42; 52.5%); (2) lymph nodes involvement (n = 22; 27.5%); (3) distant metastasis (n = 16; 20%).

Table II. Demographic and clinical characteristics of CRC patients and controls.

	Cases (*n* = 83)		Controls (*n* = 176)	
Characteristics	*n*	%	*n*	%
Demographic features				
Gender				
Male	55	66.3%	115	65.3%
Female	28	33.7%	61	34.7%
Age (years)				
Median age	71		73	
Range	28 - 91		21 - 95	
Clinical features				
Tumor location*				
Caecum	9	10.8%		
Appendix	0	0.0%		
Ascending colon	8	9.6%		
Hepatic flexure	3	3.6%		
Transverse colon	3	3.6%		
Splenic flexure	1	1.2%		
Descending colon	2	2.4%		
Sigmoid colon	18	21.7%		
Rectosigmoid junction	5	6.0%		
Rectum	30	36.1%		
Unknown	4	4.8%		
Tumor staging¥				
I	18	21.7%		
IIA – IIC	24	28.9%		
IIIA – IIIC	22	26.5%		
IVA – IVB	16	19.3%		
Unknown	3	3.6%		

*Tumor locations were defined according to the ICD-10 criteria. ¥Tumor staging was assessed according to the 7th edition of the AJCC guidelines. CRC, colorectal cancer.

Association between *IREB2* variant rs17483548 and CRC susceptibility

Aiming to evaluate the contribution of *IREB2* variant rs17483548 to CRC susceptibility, we calculated the allele and genotype frequencies of this SNP in cases and controls before comparing both groups. The genotype distributions among cases and controls did not differ significantly from those expected under HWE. However, the allele frequencies of the control group were different from those reported in genetic databases. In fact, the A allele frequency observed within the control group was 51.4% whereas the frequency reported in the 1000Genomes database to the Iberian Population in Spain is 43.9%. Nonetheless, there is

no data about *IREB2* variant rs17483548 minor allele frequency (MAF) within the Portuguese population. Also, the A allele frequency reported to the Iberian Population in Spain is the highest among all populations included in the 1000Genomes database. It is also important to note that MAF calculation includes both healthy and sick individuals, as well as persons carrying genetic variants that could protect from many diseases or increase the risk to develop them. Considering this, we recalculated the allele frequencies in all individuals enrolled in the present study (patients and controls). Therefore, the obtained allele frequencies were 52.7% for the G allele and 47.3% for the A allele, which are slightly different from the initial results and are more similar to those frequencies reported in the 1000Genomes database to the Iberian Population in Spain. Thus, we considered the A allele as the minor allele in our population.

In order to evaluate the possible association between *IREB2* variant rs17483548 and CRC, we compare both alleles in cases and controls using Fisher's exact test. The results are detailed in Table 3 and showed that A allele has a protective effect against CRC (OR = 0.593; 95% CI 0.407 - 0.864; $p = 0.006$), while G allele increases the risk to develop this neoplasm (OR = 1.687; 95% CI 1.158 – 2.457; $p = 0.006$).

The genotype frequencies of *IREB2* variant rs17483548 were also compared between CRC patients and controls by logistic regression analysis and the results are also shown in Table 3. The GG, GA and AA genotypes were found in 22.7% ($n = 40$), 51.7% ($n = 91$) and 25.6% ($n = 45$) controls, respectively. Among CRC patients, the genotype frequencies were 38.6% ($n = 32$), 45.8% ($n = 38$) and 15.7% ($n = 13$) to the same genotypes order. This means that GA is the most frequent genotype among CRC patients and controls. In turn, GG genotype is more frequent than AA genotype among CRC patients.

According to the disease association analysis, we observed that GA genotype (codominant model: OR = 0.522, 95% CI 0.287 - 0.951, $p = 0.034$) and AA genotype (codominant model: OR = 0.361, 95% CI 0.167 - 0.782, $p = 0.010$) were associated with a decreased risk for CRC development. At the same time, GG genotype (dominant model: OR = 2.133, 95% CI 1.212 - 3.755, $p = 0.009$) was found to increase the predisposition to CRC.

Table III. *IREB2* variant rs17483548 allele and genotype frequencies in cases and controls, and its association with CRC susceptibility.

	Genotype frequencies				Association analysis		
	Cases (*n* = 83)		Controls (*n* = 176)		Cases *vs* controls		
	n	%	*n*	%	OR	(95% CI)	*p*-value
Allele							
G	102	61.4%	171	48.6%	**1.687**	**(1.158 - 2.457)***	**0.006**
A	64	38.6%	181	51.4%	**0.593**	**(0.407 - 0.864)¥**	**0.006**
Genotype							
GG	32	38.6%	40	22.7%	Ref.		
GA	38	45.8%	91	51.7%	**0.522**	**(0.287 - 0.951)¥**	**0.034**
AA	13	15.7%	45	25.6%	**0.361**	**(0.167 - 0.782)¥**	**0.010**
Dominant model					**2.133**	**(1.212 - 3.755)***	**0.009**
Over-dominant model					0.789	(0.467 - 1.331)	0.374
Recessive model					0.541	(0.273 - 1.069)	0.077

The possible association between A allele (minor allele) and CRC was explored by calculating the OR (95% CI) and p-value using Fisher's exact test. To evaluate the possible associations between *IREB2* variant rs17483548 genotypes and CRC, the OR (95% CI) and p-value were calculated by logistic regression analysis according to four genetic models: codominant model (GG vs GG, GG vs GA, and GG vs AA; i.e., each genotype was compared with GG genotype of the control group as reference); dominant model (GA + AA vs GG); over-dominant model (GG + AA vs GA); recessive model (GG + GA vs AA). The values in bold indicate the existence of a statistically significant association (*susceptibility or ¥protection). G, major allele; A, minor allele; *IREB2*, Iron-responsive element binding-protein 2; CRC, colorectal cancer; OR, odds ratio; CI, confidence interval; Ref., reference.

Association between *IREB2* variant rs17483548 and CRC location and staging

In order to evaluate whether the primary location of CRC and tumor aggressiveness are influenced by *IREB2* variant rs17483548, we also assessed the possible association of this SNP with tumor location and staging at the moment of diagnosis. Tables 4 and 5 show the results of the performed statistical analysis.

Concerning to tumor location, a statistically significant result was found to the dominant model applied to evaluate the association between *IREB2* variant rs17483548 and neoplasms of the right and transverse colon, suggesting that GG genotype decreases the risk of CRC development in these locations (dominant model: OR = 0.374, 95% CI 0.162 - 0.863, $p = 0.021$). However, no other significant results were found in the association analysis between *IREB2* variant rs17483548 and all tumor locations, including left colon and rectum.

Table IV. Association between *IREB2* variant rs17483548 genotypes and CRC susceptibility for each location group.

	Cases		Association analysis		
Genotypes	*n*	%	OR	(95% CI)	*p*-value
Right and transverse colon					
GG	12	48.0%	Ref.		
GA	10	40.0%	2.380	(0.973 - 5.823)	0.058
AA	3	12.0%	3.667	(0.982 - 13.685)	0.053
Dominant model			**0.374**	**(0.162 - 0.863)¥**	**0.021**
Over-dominant model			1.552	(0.670 - 3.596)	0.305
Recessive model			2.253	(0.650 - 7.814)	0.200
Left colon					
GG	9	32.1%	Ref.		
GA	16	57.1%	1.009	(0.421 - 2.415)	0.984
AA	3	10.7%	2.619	(0.675 - 10.161)	0.164
Dominant model			0.792	(0.340 - 1.842)	0.588
Over-dominant model			0.718	(0.325 - 1.585)	0.413
Recessive model			2.604	(0.757 - 8.956)	0.129
Rectum					
GG	11	36.7%	Ref.		
GA	12	40.0%	1.758	(0.733 - 4.217)	0.206
AA	7	23.3%	1.314	(0.475 - 3.636)	0.599
Dominant model			0.627	(0.282 - 1.393)	0.252
Over-dominant model			0.638	(0.294 - 1.385)	0.256
Recessive model			0.941	(0.382 - 2.319)	0.896

The OR (95% CI) and p-value were calculated by logistic regression analysis according to four genetic models: codominant model (GG vs GG, GG vs GA, and GG vs AA; i.e., each genotype was compared with GG genotype of the control group as reference); dominant model (GA + AA vs GG); over-dominant model (GG + AA vs GA); recessive model (GG + GA vs AA). The values in bold indicate the existence of a statistically significant association (¥protection). The location groups were defined according to ICD-10 criteria: right and transverse colon (caecum, appendix, ascending colon, hepatic flexure and transverse colon); left colon (splenic flexure, descending colon, sigmoid colon and rectosigmoid junction); rectum. G, major allele; A, minor allele; *IREB2*, Iron-responsive element binding-protein 2; CRC, colorectal cancer; OR, odds ratio; CI, confidence interval; Ref., reference.

The association between *IREB2* variant rs17483548 and CRC stage was also found to be statistically significant. In fact, GA genotype (codominant model: OR = 2.227, 95% CI 1.096 - 4.526, p = 0.027) and AA genotype (codominant model: OR = 7.051, 95% CI 1.978 - 25.142, p = 0.003; recessive model: OR = 4.414, 95% CI 1.312 - 14.853, p = 0.016) were associated with an increased predisposition for locoregional CRC. On the other hand, GG genotype was found to decrease the risk of locoregional CRC (dominant model: OR = 0.347, 95% CI 0.175 - 0.685, p = 0.002). All other CRC stage groups considered in the

statistical analysis (lymph nodes involvement and distant metastasis) did not show any significant association with the *IREB2* variant rs17483548 genotypes.

Table V. Association between *IREB2* variant rs17483548 genotypes and CRC susceptibility for each tumor stage group.

	Cases		**Association analysis**		
Genotypes	*n*	%	OR	(95% CI)	*p*-value
Locoregional					
GG	20	47.6%	Ref.		
GA	19	45.2%	**2.227**	**(1.096 - 4.526)***	**0.027**
AA	3	7.1%	**7.051**	**(1.978 - 25.142)***	**0.003**
Dominant model			**0.347**	**(0.175 - 0.685)¥**	**0.002**
Over-dominant model			0.804	(0.414 - 1.560)	0.518
Recessive model			**4.414**	**(1.312 - 14.853)***	**0.016**
Lymph nodes involvement					
GG	4	18.2%	Ref.		
GA	11	50.0%	0.631	(0.193 - 2.059)	0.445
AA	7	31.8%	2.333	(0.648 - 8.401)	0.195
Dominant model			1.811	(0.591 - 5.546)	0.299
Over-dominant model			1.008	(0.421 - 2.416)	0.985
Recessive model			0.588	(0.227 - 1.518)	0.272
Distant metastasis					
GG	6	37.5%	Ref.		
GA	8	50.0%	1.375	(0.458 - 4.131)	0.570
AA	2	12.5%	2.545	(0.494 - 13.115)	0.264
Dominant model			0.621	(0.217 - 1.777)	0.375
Over-dominant model			0.992	(0.361 - 2.728)	0.987
Recessive model			2.096	(0.462 - 9.502)	0.337

The OR (95% CI) and p-value were calculated by logistic regression analysis according to four genetic models: codominant model (GG vs GG, GG vs GA, and GG vs AA; i.e., each genotype was compared with GG genotype of the control group as reference); dominant model (GA + AA vs GG); over-dominant model (GG + AA vs GA); recessive model (GG + GA vs AA). The values in bold indicate the existence of a statistically significant association (*susceptibility or ¥protection). The tumor stage groups were defined according to 7th edition of the AJCC guidelines: locoregional (stages I and IIA-IIC); lymph nodes involvement (stages IIIA-IIIC); distant metastasis (stages IVA-IVB). G, major allele; A, minor allele; *IREB2*, Iron-responsive element binding-protein 2; CRC, colorectal cancer; OR, odds ratio; CI, confidence interval; Ref., reference.

Impact on the CRC prognosis

Finally, we evaluated the impact of *IREB2* variant rs17483548 on CRC prognosis by estimating the overall survival (OS) of all patients stratified according to their genotypes, using the Kaplan-Meier method. The OS endpoint defined as deceased was observed in 11 homozygous patients for the major allele (GG genotype) and in 22 patients carrying the minor allele (GA + AA genotypes). The median survival was 24.50 months for patients with GG genotype and 18.80 months for patients with GA or AA genotypes. As mentioned above, all patients who were still alive at the moment of the last contact with the medical team were censored.

No statistically significant differences were observed in OS between patients with GG genotype and patients with GA + AA genotypes (HR = 1.269; 95% CI 0.615 - 2.620; p = 0.519). Figure 2 shows the Kaplan-Meier curves for OS in both genotype groups.

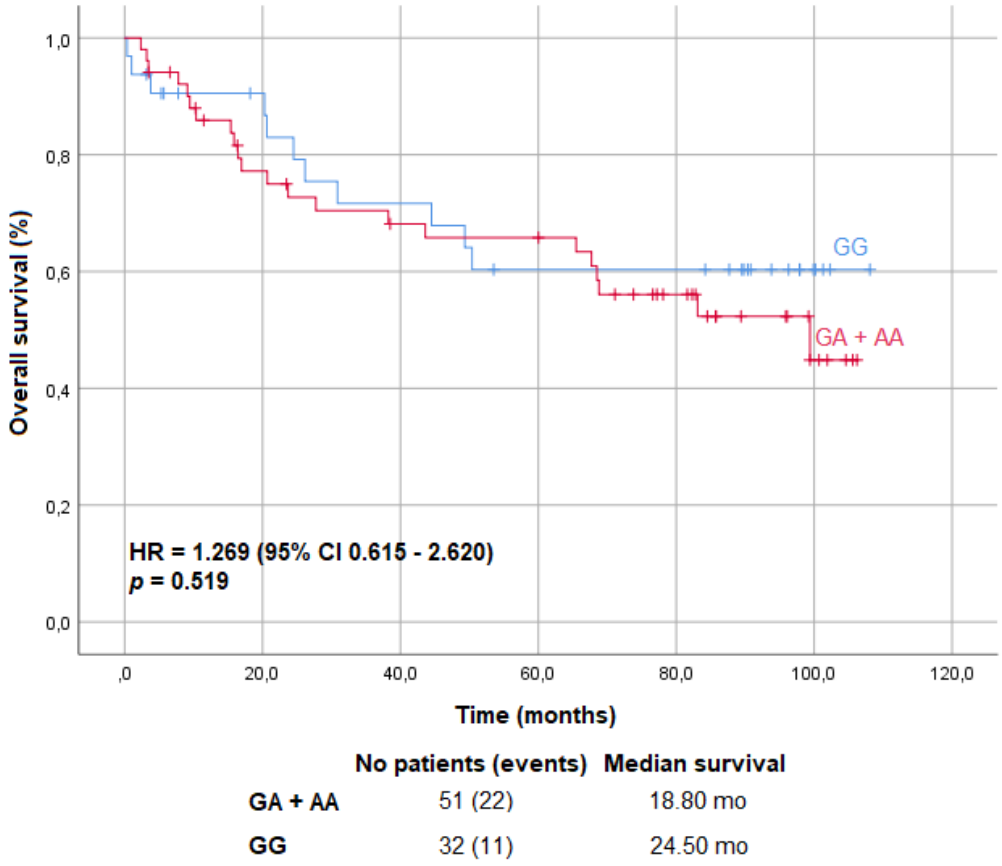

Figure 2. Overall survival curves of CRC patients according to *IREB2* variant rs17483548 genotype. The overall survival (OS) was analyzed by Kaplan-Meier method, log-rank test was used to assess differences in survival, and a multivariate Cox proportional hazards model was performed to estimate hazard ratio (HR) and its 95% confidence interval (CI). No, number; mo, months; GG, homozygous patients for the G allele; GA, heterozygous patients, AA, homozygous patients for the A allele.

DISCUSSION AND CONCLUSION

The genetic background and variability has an important role in CRC predisposition, initiation, and progression.[5,29] Some hereditary forms, such as Lynch syndrome and familial adenomatous polyposis, are caused by well-known mutations with high penetrance but account for less than 5% of all cases. However, the entire genetic contribution for CRC development could be responsible for up to 35% of all cases.[5,29,30]

In recent years, numerous genome-wide association studies (GWAS) and case-control studies have identified a growing number of susceptibility loci and SNPs associated with an increased risk for CRC development. To date, approximately 70 susceptibility variants from 42 genetic regions were identified using GWAS.[31] Individually, these susceptibility variants account only for a modest risk, but a large proportion of the population carries one or more risk alleles because of the high allele frequencies of these susceptibility variants.[29,30] The cumulative effect of carrying multiple risk alleles and its integration with family history and other various environmental and non-genetic factors could be useful to develop new screening tests and diagnostic algorithms. This concept has been studied for many cancers, including CRC, with some positive and promising results.[30,32–34]

However, the individualized CRC risk profiling has still poor performances because the majority of susceptibility SNPs was not yet discovered. The continuous effort to discover novel susceptibility loci and common genetic variants associated with an increased or decreased risk for CRC development is extremely important to improve the effectiveness of these screening approaches. Study the complex interaction between genetic variants and environmental exposures is also necessary to improve our understanding of all biological pathways and mechanisms underlying the CRC carcinogenic process. With the integration of these insights, CRC risk profiling would expectably play a pivotal role in CRC prevention.

In the present study, we assessed the influence of the *IREB2* variant rs17483548 on the susceptibility to CRC, by performing a hospital-based case-control study. Furthermore, we also evaluated the potential associations between this SNP and CRC location, stage, and prognosis. The observed differences in the allele frequencies between the control group enrolled in the present study and the reported frequencies in genetic databases can be explained by the global diversity of human genomes, which is the outcome of a wide variety of demographic and evolutionary events (such as migration, population isolation, admixture, bottlenecks, genetic drift, and natural selection) occurred in different parts of the world at

different time points.[35] However, considering all individuals enrolled in the present study, we obtained an A allele frequency of 47.3% which are very similar to the MAF (referring to A allele frequency) reported in the 1000Genomes database for the Iberian Population in Spain (43.9%).

The association analysis showed that patients carrying the G allele have an increased susceptibility to CRC development. In contrast, the A allele is a protective factor against CRC. These results were corroborated by the association analysis performed to assess the relationship between *IREB2* variant rs17483548 genotypes and CRC susceptibility. In fact, we observed that patients carrying the GA and the AA genotypes have approximately 2-fold and 3-fold lower susceptibility to CRC development, respectively. On the other hand, homozygous patients for the G allele have a 2-fold increase in the CRC risk. To our knowledge, it was the first time that *IREB2* variant rs17483548 was associated with cancer. Previous studies with this SNP showed statistically significant associations with age-related macular degeneration[36] and nicotine dependence[37], but no published studies assessed its association with CRC or other cancer types.

Concerning to CRC location, only the GG genotype has a statistically significant association with CRC of the right and transverse colon. Patients carrying GG genotype have a lower risk to develop CRC in the mentioned locations. However, the association of GA (p = 0.058) and AA (p = 0.053) genotypes with CRC of the right and transverse colon almost reached the statistical significance too. The OR for these latter associations suggests that patients carrying GA genotype (OR = 2.380) and AA genotype (OR = 3.667) may have an increased susceptibility to develop CRC in the right and transverse colon. Using a larger sample size, it could be possible to determine if these associations are statistically significant. Clarify this information is important because CRC is a heterogeneous disease highly affected by the anatomical location and microenvironment of the primary tumor. The right-sided tumors tend to be more difficult to diagnose than left-sided tumors due to their flat morphology. For this reason, right-sided tumors are commonly diagnosed in more advanced stages than left-sided tumors.[38] Know the genetic variants more associated with each tumor location could be useful to stratify the risk for right-sided and for left-sided tumors, allowing to identify the patients who could benefit with more intensive screening protocols. Furthermore, right-sided CRC has a poorer prognosis when compared with tumors located in the left colon and responds differently to the non-surgical treatments.[39] Understand the molecular and genetic factors for these differences is extremely important to adequate treatment strategies.

Therefore, the evaluation of the *IREB2* variant rs17483548 influence in CRC primary location may be of great relevance in further studies.

The association analysis performed to evaluate the relationship between *IREB2* variant rs17483548 genotypes and CRC staging showed that GA and AA genotypes are risk factors for locoregional cancer, while GG genotype decreases the susceptibility for this early stage cancer. The codominant and recessive models demonstrated that AA genotype is associated with the higher risk. Despite being risk genotypes, the obtained results could indicate that GA and AA genotypes are associated with low aggressive tumors since locoregional stage is the lowest considered in this study. In addition, no other relevant associations between these both genotypes and more advanced tumor stages were observed. For the same principle, patients carrying GG genotype have a lower susceptibility to locoregional CRC but could be more prone to develop advanced disease. However, this possibility was not supported by the results because no statistically significant associations were observed with lymph nodes involvement and metastatic disease stages.

Finally, the overall survival analysis did not show any impact of *IREB2* variant rs17483548 on the CRC patients prognosis. However, the median survival of the patients carrying GA and AA genotypes, which were found to be protective genotypes against CRC development, was paradoxically lower than the median survival of the homozygous patients for the G allele.

Some important limitations should be considered while interpreting the results of the present study. First, all patients and controls enrolled in this research were recruited from a single hospital of the central region of Portugal. This fact renders it impossible to determine if the obtained results can be extended to the general population and to other regions of the world. Second, the sample size is relatively small which may limit the statistical power. However, the allele frequencies observed in our patients and controls groups were in HWE, which suggests that our subject sampling was sufficiently random and no selection bias was introduced. Third, we did not take into account other factors known to influence CRC risk, such as family history of CRC, smoking, consumption of alcohol, consumption of red and processed meat, physical activity, body mass index, diabetes, hormone replacement therapy, or aspirin use. These factors have an important impact in CRC development and need to be considered in further studies to avoid confounding bias. Furthermore, the interaction of *IREB2* variant rs17483548 with these factors should also be studied, because genome-environment interactions are key determinants of CRC risk. Finally, the functional impact of *IREB2* variant rs17483548 was not explored in this study. Nevertheless, this

information is extremely important to understand the molecular mechanisms through which this variant influences CRC risk and carcinogenic process.

Further studies are needed to characterize better the impact of this SNP in CRC susceptibility, location and staging. Considering the fundamental role of IRP2 in the regulation of iron metabolism, it is also important to assess the impact of *IREB2* variant rs17483548 in iron homeostasis since the available information is still scarce.

Altogether, the data presented in this paper should be interpreted as preliminary evidence and must be replicated by other studies. Nevertheless, it suggests that *IREB2* variant rs17483548 may play an important role in CRC susceptibility and staging.

AKNOLEDGEMENTS

The execution of the present work would not be possible without the tireless help and the unconditional support of several people. For this reason, I would like to thank all those who supported me in this long journey.

To my parents, Joaquim and Susana, I thank the encouragement and the support they give to me every day with the same enthusiasm of who share with me the same dreams and ambitions.

To my brother, Bruno, I thank his keen interest in all my challenges and for being my best friend in every occasion. He is the person who I can count whenever I need and who wholeheartedly supports me.

To my maternal grandparents, Inês and Abel, I thank the affectionate way they always helped me and encouraged me during my all education, but above all for the values and guidance they gave me since childhood which made me the person I am today.

To my grandmother Gloria, I thank her kindness and the tender way she encourages me to improve every day, reminding me the importance of being a good man before wishing to be a good doctor.

To my girlfriend, Tatiana, I thank her presence in all the good and less good moments of this long 6-years walk. Her support and her contagious and inexhaustible energy were also an indispensable help along this work. I wish to continue counting with this energy and this support in new and greater challenges.

To Dr. Joana Jorge and Doctor Raquel Alves, I thank all lessons since my first day in the laboratory and all indispensable help they gave to me in the execution of all laboratory procedures. I am sure that this work would not have been possible without the inexhaustible knowledge, availability and patience which they always provided to me.

To Professor Doctor Ana Cristina Gonçalves, I thank the tireless support she gave me in all the stages of this work. Her help with the laboratory procedures and with the statistical analysis was decisive for the completion of this work, as well as the knowledge, the

competence, and the recommendations she transmitted to me. I certainly will take this knowledge with me along my professional career.

To my co-advisor, Professor Doctor Ana Bela Sarmento-Ribeiro, I thank having welcomed me on her research team and all the commitment and knowledge she offered to me along this work. Her orientation and readiness were fundamental for the execution of all steps of this work.

To my advisor, Professor Doctor José Nascimento Costa, I thank all the enthusiasm, confidence and knowledge he always transmitted to me along the execution of this work. I also would like to thank his indispensable guidance and readiness whenever necessary without which this work would not have been possible.

To my friends Diana Andrade, João Lima, and Duarte Gil, I thank their presence in every step of this work and for encouraging me and helping me whenever I needed it. I am sure that all human qualities which differentiate them will also make them excellent doctors. I hope we can continue to share our future conquests.

CONFLICT OF INTEREST

The authors do not have any conflict of interest to declare.

AGRADECIMENTOS

A realização deste trabalho não teria sido possível sem a ajuda incansável e o apoio incondicional de várias pessoas. Gostaria, por isso, de deixar uma palavra de agradecimento a todos os que me apoiaram nesta longa caminhada.

Aos meus pais, Joaquim e Susana, agradeço o incentivo que me dão todos os dias e o apoio que sempre me ofereceram com o entusiasmo de quem partilha comigo os mesmos sonhos e ambições.

Ao meu irmão, Bruno, agradeço o enorme interesse que sempre demonstrou pelos desafios que tenho enfrentado e por ser o amigo de todas as ocasiões com quem posso contar sempre que preciso e que me tem apoiado incondicionalmente.

Aos meus avós maternos, Inês e Abel, agradeço a forma carinhosa com que sempre me ajudaram e incentivaram ao longo de todo o curso, mas sobretudo pelos valores e ensinamentos que me transmitiram desde criança e que fazem de mim a pessoa que sou hoje.

À minha avó Glória agradeço a bondade que sempre me transmitiu e a forma enternecedora com que me incentiva a melhorar todos os dias, lembrando-me da importância de ser um bom homem antes de pretender ser um bom médico.

À minha namorada, Tatiana, agradeço o facto de ter estado presente em todos os momentos bons e menos bons desta longa caminhada de 6 anos. O seu apoio e a sua energia contagiante e inesgotável foram uma ajuda imprescindível em todas as etapas deste trabalho. Desejo contar com eles em novos e maiores desafios.

À Dra. Joana Jorge e à Dra. Raquel Alves agradeço todos os ensinamentos que me transmitiram desde o meu primeiro dia no laboratório e toda a ajuda indispensável que me prestaram na execução dos procedimentos laboratoriais. Estou certo que não teria sido possível concretizar este trabalho sem a sabedoria, a disponibilidade e a paciência inesgotáveis que sempre me dispensaram.

À Professora Doutora Ana Cristina Gonçalves, agradeço o apoio incansável que me deu em todas as etapas da realização deste trabalho. A sua ajuda nos procedimentos laboratoriais e na análise estatística foram determinantes para a conclusão deste trabalho, assim como os conhecimentos, as competências e os conselhos que me transmitiu e que certamente me acompanharão ao longo da minha carreira profissional.

À minha coorientadora, Professora Doutora Ana Bela Sarmento-Ribeiro, agradeço por ter-me aceite na sua equipa de investigação e por todo o empenho e sabedoria que me dispensou ao longo da realização deste trabalho. A sua orientação e disponibilidade foram fundamentais para a execução de todas as etapas deste trabalho.

Ao meu orientador, Professor Doutor José Nascimento Costa, agradeço o entusiasmo, a confiança e a sabedoria que sempre me transmitiu na realização deste trabalho. Agradeço também a orientação e a disponibilidade imprescindíveis que me dispensou sempre que necessário e sem as quais não teria sido possível concretizar este trabalho.

Aos meus amigos Diana Andrade, João Lima e Duarte Gil, deixo uma palavra de enorme agradecimento por terem estado presentes em todos os momentos da execução deste trabalho e por me incentivarem e ajudarem sempre que necessitei. Estou certo que as qualidades humanas que vos diferenciam farão de vós excelentes médicos. Espero poder continuar a partilhar convosco as conquistas que formos alcançando.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA - A Cancer J Clin. 2018;68(6):394–424.

2. Colorectal cancer [Internet]. Globocan 2018. 2019 [cited 2019 Mar 26]. Available from: https://gco.iarc.fr/today/data/factsheets/cancers/10_8_9-Colorectum-fact-sheet.pdf

3. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683–91.

4. Yang T, Li X, Montazeri Z, Little J, Farrington SM, Ioannidis JPA, et al. Gene-environment interactions and colorectal cancer risk: an umbrella review of systematic reviews and meta-analyses of observational studies. Int J Cancer. 2018;doi: 10.1002/ijc.32057 [Epub ahead of print].

5. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet. 2014;383(9927):1490–502.

6. Brenner H, Chen C. The colorectal cancer epidemic: challenges and opportunities for primary, secondary and tertiary prevention. Br J Cancer. 2018;119(7):785–792.

7. Nelson RL. Dietary iron and colorectal cancer risk. Free Radic Biol Med. 1992;12(2):161–8.

8. Nelson RL, Davis FG, Sutter E, Sobin LH, Kikendall JW, Bowen P. Body iron stores and risk of colonic neoplasia. J Natl Cancer Inst. 1994;86(6):455–60.

9. Wurzelmann JI, Silver A, Schreinemachers DM, Sandler RS, Everson RB. Iron intake and the risk of colorectal cancer. Cancer Epidemiol Biomarkers Prev. 1996;5(7):503–7.

10. Chua ACG, Klopcic B, Lawrance IC, Olynyk JK, Trinder D. Iron: An emerging factor in colorectal carcinogenesis. World J Gastroenterol. 2010;16(6):663–72.

11. Cross AJ, Ferrucci LM, Risch A, Graubard BI, Ward MH, Park Y. A large prospective study of meat consumption and colorectal cancer risk: an investigation of potential mechanisms underlying this association. Cancer Res. 2010;70(6):2406–14.

12. Torti SV, Torti FM. Iron and cancer: more ore to be mined. Nat Rev Cancer.

2013;13(5):342–55.

13. Dizdaroglu M, Jaruga P. Mechanisms of free radical-induced damage to DNA. Free Radic Res. 2012;46(4):382–419.

14. Manz DH, Blanchette NL, Paul BT, Torti FM, Torti SV. Iron and cancer: recent insights. Ann N Y Acad Sci. 2016;1368(1):149–61.

15. Pfeifhofer-Obermair C, Tymoszuk P, Petzer V, Weiss G, Nairz M. Iron in the Tumor Microenvironment - Connecting the Dots. Front Oncol. 2018;8:549.

16. Guo B, Brown FM, Phillips JD, Yu Y, Leibold EA. Characterization and expression of iron regulatory protein 2 (IRP2). Presence of multiple IRP2 transcripts regulated by intracellular iron levels. J Biol Chem. 1995;270(28):16529–35.

17. LaVaute T, Smith S, Cooperman S, Iwai K, Land W, Meyron-Holtz E, et al. Targeted deletion of the gene encoding iron regulatory protein-2 causes misregulation of iron metabolism and neurodegenerative disease in mice. Nat Genet. 2001;27(2):209–14.

18. Pantopoulos K. Iron metabolism and the IRE/IRP regulatory system: an update. Ann N Y Acad Sci. 2004;1012:1–13.

19. Wang W, Di X, D'Agostino RBJ, Torti SV, Torti FM. Excess capacity of the iron regulatory protein system. J Biol Chem. 2007;282(34):24650–9.

20. Horniblow RD, Bedford M, Hollingworth R, Evans S, Sutton E, Lal N, et al. BRAF mutations are associated with increased iron regulatory protein-2 expression in colorectal tumorigenesis. Cancer Sci. 2017;108(6):1135–43.

21. Hamara K, Bielecka-Kowalska A, Przybylowska-Sygut K, Sygut A, Dziki A, Szemraj J. Alterations in expression profile of iron-related genes in colorectal cancer. Mol Biol Rep. 2013;40(10):5573–85.

22. Maffettone C, Chen G, Drozdov I, Ouzounis C, Pantopoulos K. Tumorigenic Properties of Iron Regulatory Protein 2 (IRP2) Mediated by Its Specific 73-Amino Acids Insert. PLoS One. 2010;5(4):e10163.

23. Rychtarcikova Z, Lettlova S, Tomkova V, Korenkova V, Langerova L, Simonova E, et al. Tumor-initiating cells of breast and prostate origin show alterations in the expression of genes related to iron metabolism. Oncotarget. 2017;8(4):6376–98.

24. Deng Z, Manz DH, Torti SV, Torti FM. Iron-responsive element-binding protein 2 plays an essential role in regulating prostate cancer cell growth. Oncotarget. 2017;8(47):82231–43.

25. Khiroya H, Moore JS, Ahmad N, Kay J, Woolnough K, Langman G, et al. IRP2 as a potential modulator of cell proliferation, apoptosis and prognosis in nonsmall cell lung cancer. Eur Respir J. 2017;49(4):pii: 1600711.

26. Wang W, Deng Z, Hatcher H, Miller LD, Di X, Tesfay L, et al. IRP2 regulates breast tumor growth. Cancer Res. 2014;74(2):497–507.

27. Prokunina L, Alarcón-Riquelme ME. Regulatory SNPs in complex diseases: their identification and functional validation. Expert Rev Mol Med. 2004;6(10):1–15.

28. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215.

29. de la Chapelle A. Genetic predisposition to colorectal cancer. Nat Rev Cancer. 2004;4(10):769–80.

30. Dunlop MG, Tenesa A, Farrington SM, Ballereau S, Brewster DH, Koessler T, et al. Cumulative impact of common genetic variants and other risk factors on colorectal cancer risk in 42,103 individuals. Gut. 2013;62(6):871–81.

31. Schmit SL, Edlund CK, Schumacher FR, Gong J, Harrison TA, Huyghe JR, et al. Novel Common Genetic Susceptibility Loci for Colorectal Cancer. J Natl Cancer Inst. 2019;111(2):146–57.

32. Hsu L, Jeon J, Brenner H, Gruber SB, Schoen RE, Berndt SI, et al. A model to determine colorectal cancer risk using common genetic susceptibility loci. Gastroenterology. 2015;148(7):1330–1339.e14.

33. Weigl K, Thomsen H, Balavarca Y, Hellwege JN, Shrubsole MJ, Brenner H. Genetic Risk Score is Associated With Prevalence of Advanced Neoplasms in a Colorectal Cancer Screening Population. Gastroenterology. 2018;155(1):88–98.e10.

34. Jenkins MA, Makalic E, Dowty JG, Schmidt DF, Dite GS, MacInnis RJ, et al. Quantifying the utility of single nucleotide polymorphisms to guide colorectal cancer screening. Futur Oncol. 2016;12(4):503–13.

35. Choudhury A, Hazelhurst S, Meintjes A, Achinike-Oduaran O, Aron S, Gamieldien J,

et al. Population-specific common SNPs reflect demographic histories and highlight regions of genomic plasticity with functional relevance. BMC Genomics. 2014;15:437.

36. Synowiec E, Pogorzelska M, Blasiak J, Szaflik J, Szaflik JP. Genetic polymorphism of the iron-regulatory protein-1 and -2 genes in age-related macular degeneration. Mol Biol Rep. 2012;39(6):7077–87.

37. Culverhouse RC, Saccone NL, Stitzel JA, Wang JC, Steinbach JH, Goate AM, et al. Uncovering hidden variance: pair-wise SNP analysis accounts for additional variance in nicotine dependence. Hum Genet. 2011;129(2):177–88.

38. Baran B, Ozupek NM, Tetik NY, Acar E, Bekcioglu O, Baskin Y. Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. Gastroenterol Res. 2018;11(4):264–73.

39. Stintzing S, Tejpar S, Gibbs P, Thiebach L, Lenz HJ. Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. Eur J Cancer. 2017;84:69–80.