

**ANÁLISE DA SUSCETIBILIDADE AO CÂNCER
COLORRETAL EM INDIVÍDUOS BRASILEIROS
POR GENOTIPAGEM DE POLIMORFISMOS DE
NUCLEOTÍDEO ÚNICO CONHECIDOS**

ISRAEL GOMY

**Tese apresentada à Fundação Antônio
Prudente para obtenção do título de Doutor
em Ciências**

Área de concentração: Oncologia

**Orientador: Prof. Dr. Benedito Mauro Rossi
Co-Orientadora: Prof^a Sílvia Regina Rogatto**

**São Paulo
2013**

FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Gomy, Israel

Análise da suscetibilidade ao câncer colorretal em indivíduos brasileiros por genotipagem de polimorfismos de nucleotídeo único conhecidos / Israel Gomy – São Paulo, 2013.

82p.

Tese (Doutorado)-Fundação Antônio Prudente.

Curso de Pós-Graduação em Ciências - Área de concentração:

Oncologia.

Orientador: Benedito Mauro Rossi

Descritores: 1. NEOPLASIAS COLORRETAIS. 2. TÉCNICAS DE GENOTIPAGEM. 3. POLIMORFISMO GENÉTICO.

"If there were no individual variability, Medicine would have been Science not an Art."

Sir William Osler (1849-1919)

DEDICATÓRIA

À minha família (pais, irmã, sobrinhas) e aos meus afilhados.

AGRADECIMENTOS

Ao meu orientador, Prof^o Dr. Benedito Mauro Rossi e co-orientadora Prof^a Sílvia Regina Rogatto.

À equipe de Coloproctologia do Hospital A.C. Camargo, especialmente ao Prof. Dr. Fábio de Oliveira Ferreira, Dr. Alexander Bressan, Dr. Wilson Nakagawa e Dr. Samuel Aguiar Júnior.

À equipe de Oncologia Clínica do Hospital A.C. Camargo, especialmente ao Dr. Marcelo Fanelli e Dr. Celso Abdon Mello.

À equipe de Medicina Comunitária do Hospital A.C. Camargo, enfermeiros e técnicos de enfermagem do laboratório de coleta do Hospital A.C. Camargo.

À equipe de Enfermagem da Coloproctologia do Hospital A.C. Camargo, especialmente à Érika Monteiro dos Santos e suas alunas.

À equipe do Banco de Macromoléculas do Hospital A.C. Camargo, especialmente à Louise Danielle de Carvalho Mota e Eloisa Olivieri.

À equipe do Laboratório Neogene do Centro de Pesquisas do Hospital A.C. Camargo, especialmente à Graziela, Sandra, Sandro, Deborah, Ana, Márcia e Tatiana.

À equipe do Laboratório de Genômica e Biologia Molecular do Centro de Pesquisas do Hospital A.C. Camargo, especialmente à Dra. Dirce Carraro.

À equipe do Serviço de Arquivos Médicos do Hospital A.C. Camargo.

À equipe da Biblioteca do Hospital A.C. Camargo.

À equipe da Pós-graduação do Hospital A.C. Camargo.

À Fundação Antônio Prudente, FAPESP, CAPES e CNPq.

RESUMO

Gomy I. **Análise da suscetibilidade ao câncer colorretal em indivíduos brasileiros por genotipagem de polimorfismos de nucleotídeo único conhecidos.** São Paulo; 2013. [Tese de Doutorado-Fundação Antônio Prudente]

O câncer colorretal (CCR) é um dos cânceres mais prevalentes no mundo, o terceiro mais comum em homens e o segundo em mulheres, sendo a quarta maior causa de mortes por câncer. As taxas de incidência do CCR vêm aumentando nos países em desenvolvimento, como o Brasil, representando a terceira neoplasia maligna mais frequente em ambos os sexos. Suas causas são multifatoriais e interdependentes, com fatores de risco genéticos e ambientais, sendo a história familiar um dos principais componentes. A herdabilidade do CCR é de aproximadamente 35%, sendo que as síndromes hereditárias respondem por até 6% dos casos. Recentemente, estudos de associação em larga escala genômica (GWAS) vêm demonstrando que parte do risco genético se deve a variantes comuns de baixa penetrância, como os polimorfismos de nucleotídeo único (*SNPs*). Até o momento, GWAS em populações de ancestralidade europeia identificaram cerca de 20 *SNPs* que, individualmente, conferem pouco efeito no aumento do risco do CCR mas, conjuntamente, respondem por grande parte do risco populacional. Estudos replicativos em populações não-europeias têm obtido pouco poder para detectar associações significativas. Diante disso, os objetivos do presente estudo foram: identificar dez *SNPs* descritos em populações europeias em indivíduos da população brasileira; calcular as frequências alélica e genotípica dos dez *SNPs* nos casos e controles; determinar a magnitude do efeito dos alelos sobre o risco de CCR e correlacionar os alelos de risco com as características clínico-patológicas e a história familiar. Foram incluídos 1.467 indivíduos (727 casos e 740 controles) que, após assinatura do TCLE, coletaram sangue para a amplificação do DNA e

genotipagem dos *SNPs* na plataforma de qPCR 7900HT com ensaios *TaqMan* e análise com o software SDS2.3. A taxa global de genotipagem foi maior do que 95% e não houve desequilíbrio de Hardy-Weinberg entre os casos e controles. 51% dos casos eram do sexo masculino, com média da idade ao diagnóstico de 57 anos; 30% preencheram os critérios de Bethesda; 3% dos tumores foram adenomas avançados; a localização mais comum foi o reto e o estadiamento III foi o mais frequente ao diagnóstico, com a maior parte dos tumores invadindo estruturas pericólicas/perirretais, porém, sem metástases à distância; predominaram os tumores moderadamente diferenciados e com tipo histológico tubular. A maior parte dos pacientes estava viva sem doença e cerca de 30% dos casos não apresentava história familiar de CCR em parentes de até 3º grau; 52% dos controles eram do sexo feminino, com média de idade de 52 anos. Dos dez *SNPs*, metade (06, 09 16, 82, 83) apresentou associação com significância corrigida por múltiplos testes na maioria dos modelos genéticos, e dois *SNPs* (26 e 71) tenderam à associação. O *SNP* 06 obteve a associação mais significativa, com maior poder estatístico e maior efeito sobre o risco. O *SNP* 09 foi o único que apontou proteção ao risco. A falta de associação entre os *SNPs* 21, 45, 59, provavelmente deveu-se à falta de poder pelo tamanho insuficiente da amostra. Não foi realizada a correção de possíveis falsos positivos pela estratificação por ancestralidade, necessária em populações miscigenadas como a brasileira. No entanto, as frequências dos alelos de risco não diferiram significativamente com as das populações européias, e os efeitos foram similarmente pequenos. Entre as correlações genótipo-fenótipo, o *SNP* 06 apontou como fator de bom prognóstico, e o *SNP* 82 com os critérios revisados de Bethesda. Entretanto, a aplicabilidade desses fatores na prática clínica é limitada pela falta de estudos prospectivos. Este estudo replicou parcialmente os GWAS europeus e reforçou a necessidade de estratificar a população brasileira.

SUMMARY

Gomy I. **[Analysis of the susceptibility to colorectal cancer in Brazilian Individuals through genotyping known single nucleotide polymorphisms]**. São Paulo; 2013. [Tese de Doutorado-Fundação Antônio Prudente]

Colorectal cancer (CRC) is one of the most prevalent cancers worldwide, the third most common cancer in men and the second one in women, and the fourth leading cause of all cancer deaths. CRC incidence rates have been increasing specially in developing countries, such as Brazil, where is the third most frequent cancer in both genders. Etiology of CRC is multifactorial, both environmental and genetic risk factors interacting with each other, of whom family history plays a special role. CRC heritability is about 35% and Mendelian syndromes respond for 6% of the cases. Recently, genome-wide association studies (GWAS) have showed that part of the risk is due to common low penetrant variants, such as single nucleotide polymorphisms (SNPs). Approximately 20 SNPs have been discovered through GWAS from European-descent populations, each one with modest size effects on the CRC risk, but collectively, make great part of populational risk. Non-European replication studies have not achieved enough power to detect significant association. Thus, the aims of the present study were: (1) to identify ten SNPs previously detected in European populations in the Brazilian population; (2) to calculate allelic and genotypic frequencies of the ten SNPs in cases and controls; (3) to detect effect sizes of the risk alleles and (4) correlate risk with clinical pathological characteristics and family history. 1,467 individuals (727 cases and 740 controls) were included, who signed consent forms to get their blood drawn for DNA amplification and SNP genotyping through qPCR in 7900HT platform with *TaqMan* assays and analysis with SDS2.3 software. The overall genotyping rate was above 95% and there was no deviation from Hardy-Weinberg equilibrium among cases

and controls. 51% of cases were male, with mean age at diagnosis of 57 years-old; 30% fulfilled Bethesda criteria; 3% were advanced adenomas; rectal and stage III tumors were most frequent at diagnosis, with pericolic/perirectal invasion and without distant metastasis; grade 2 differentiated tubular tumors predominated. The majority of patients were alive and healthy and about one third had no CRC family history; 52% of controls were female with mean age of 52 years-old. Half of the ten SNPs (06, 09, 16, 82, 83) significantly associated with CRC risk after correction for multiple tests in most genetic models, whereas two tended to be associated (26 e 71). SNP 06 had the most significant association, strongest statistical power and greater effect size on CRC risk. SNP 09 was the only one to be protective. Lack of association among SNPs 21, 45, 59 was most likely due to insufficient power by the small sample size. Correction for eventual false positives through ancestry stratification was not performed, although it is necessary for admixed populations. However, risk allele frequencies did not significantly differ from European as their effects were similarly small. Genotype-phenotype correlations showed SNP 06 as a good prognostic factor and SNP 82 as associated with revised Bethesda criteria. Nevertheless, clinical application of these factors is limited by the lack of prospective studies. This study partially replicated European GWAS and reinforced the need to stratify the Brazilian population.

LISTA DE FIGURAS

Figura 1	As três classes de alelos de suscetibilidade ao câncer.....	11
Figura 2	A distribuição dos alelos de risco em casos e controles.....	18
Figura 3	Representação esquemática da genotipagem de um indivíduo heterozigoto.....	28
Figura 4	Representação esquemática da genotipagem de um indivíduo homozigoto 1.....	28
Figura 5	Representação esquemática da genotipagem de um indivíduo homozigoto 2.....	28
Figura 6	Exemplo de <i>clusters</i> satisfatórios das amostras.....	29
Figura 7	Exemplo de <i>clusters</i> insatisfatórios das amostras.....	30
Figura 8	Cálculos do poder estatístico.....	47

LISTA DE TABELAS

Tabela 1	Dados dos GWAS em CCR com os <i>SNPs</i> identificados até o momento.....	17
Tabela 2	Características clínico-patológicas dos casos.....	36
Tabela 3	Frequências genótípicas em casos e controles.....	38
Tabela 4	Frequências alélicas nos casos e controles.....	39
Tabela 5	Testes de associação genética.....	40
Tabela 6	Os cinco <i>SNPs</i> associados com o risco de CCR.....	42
Tabela 7	Correlações genótipo-fenótipo.....	43
Tabela 8	Dados dos 10 <i>SNPs</i> identificados pelos GWAS em populações européias.....	56

LISTA DE SIGLAS E ABREVIATURAS

BMP	<i>Bone Morphogenetic Protein</i>
CCR	Câncer colorretal
DL	Desequilíbrio de ligação
EHW	Equilíbrio de Hardy-Weinberg
F_{Am}	Frequência do alelo <i>minor</i>
GWAS	<i>Genome-wide Association Studies</i>
HACC	Hospital do Câncer A.C.Camargo
IC	Intervalo de confiança
INCA	Instituto Nacional de Câncer
MIA	Marcadores informativos de ancestralidade
OR	<i>Odds Ratio</i>
qPCR	Reação em cadeia da polimerase em tempo real
SNPs	<i>Single Nucleotide Polymorphisms</i>
TCLE	Termo de consentimento livre e esclarecido
TGF-β	<i>Transforming Growth Factor β</i>

ÍNDICE

1	INTRODUÇÃO	1
1.1	Epidemiologia	1
1.2	Etiopatogênese.....	3
1.3	Polimorfismos de nucleotídeo único e suscetibilidade genética	6
1.4	Suscetibilidade genética ao câncer colorretal.....	9
1.5	Estudos de associação de larga escala genômica.....	13
1.5.1	Estudos de associação de larga escala genômica e câncer colorretal.....	16
1.6	Implicações futuras.....	20
2	OBJETIVOS	22
3	MATERIAL E MÉTODOS.....	23
3.1	Casuística.....	23
3.2	Metodologia	25
3.2.1	Coleta e processamento das amostras	25
3.2.2	Amplificação do DNA.....	26
3.2.3	Genotipagem dos <i>SNPs</i>	26
3.3	Análises estatísticas	30
4	RESULTADOS	35
4.1	Características dos casos e controles	35
4.2	Genotipagem de <i>SNPs</i> e testes de associação.....	37
4.3	Correlações genótipo-fenótipo.....	43
4.4	Tamanho da amostra e poder estatístico	46
5	DISCUSSÃO	51
5.1	Poder estatístico	51
5.2	Frequências e efeitos dos <i>SNPs</i>	55
5.3	Estratificação por ancestralidade.....	57

5.4	Correlações genótipo-fenótipo.....	63
5.5	Considerações finais	70
6	CONCLUSÃO	72
7	REFERÊNCIAS BIBLIOGRÁFICAS	73

ANEXOS

- Anexo 1** Termo de Consentimento Livre e Esclarecido para Coleta de Sangue (Casos)
- Anexo 2** Termo de Consentimento Livre e Esclarecido para Coleta de Sangue (Controles)
- Anexo 3** Termo de Consentimento Livre e Esclarecido para Amostras do Banco de Macromoléculas (Casos)
- Anexo 4** Ficha para Coleta de Dados Clínico-Patológicos e História Familiar

1 INTRODUÇÃO

1.1 EPIDEMIOLOGIA

O câncer colorretal (CCR) é um dos cânceres mais prevalentes no mundo, responsável por cerca de 10% de todas as neoplasias malignas. Estima-se que a incidência anual global ultrapasse um milhão de casos novos, representando o terceiro câncer mais comum em homens (663.000 casos) e o segundo em mulheres (570.000 casos), e a quarta maior causa de mortes por câncer (603.000), o que corresponde a 8% de todas as mortes por câncer no mundo (FERLAY et al. 2010).

A variação nas taxas de incidência global do CCR está diminuindo progressivamente, e o aumento de novos casos em populações historicamente com baixa incidência, como a asiática, parece refletir um maior desenvolvimento econômico e, conseqüentemente, uma maior exposição aos fatores de risco denominados “estilo de vida ocidentalizado”. As maiores taxas de incidência estão em países como a Austrália e Nova Zelândia e na Europa Ocidental, mas o aumento é maior em países da Europa Oriental, Israel e Japão, enquanto que se estabilizou na Europa Setentrional e declinou nos Estados Unidos, provavelmente pela detecção e remoção precoces de lesões pré-malignas (EDWARDS et al. 2010). As taxas de mortalidade do CCR, por sua vez, vêm diminuindo na maior parte dos países desenvolvidos, provavelmente devido aos programas de

rastreamento e às novas modalidades terapêuticas. Entretanto, observa-se um aumento em países em desenvolvimento, como a Rússia, México, Chile e Brasil, refletindo, possivelmente, o estilo de vida e o aumento da expectativa de vida, assim como de melhores registros de dados e a carência de políticas públicas de rastreamento populacional. Dados de registros internacionais apontam que a taxa de sobrevida global em cinco anos é, em média, de 55% nos países desenvolvidos e 40% nos países em desenvolvimento e é maior nas mulheres do que nos homens (CENTER et al. 2009).

Segundo estimativas do Instituto Nacional de Câncer para o biênio 2012/2013, no Brasil, 30.140 indivíduos desenvolverão CCR (15.960 mulheres e 14.180 homens), correspondendo à terceira neoplasia maligna mais frequente em ambos os sexos, com um risco estimado de 15 casos novos a cada 100.000 homens e 16 a cada 100.000 mulheres. Já na região Sudeste do país, o CCR é o segundo mais frequente em ambos os sexos (22/100.000 homens e 23/100.000 mulheres). A razão de incidência homem/mulher será de 0,88, enquanto que a razão global é de 1,4. As taxas de mortalidade, por sua vez, foram 5,9/100.000 em homens e 5,6/100.000 em mulheres em 2004 (Ministério da Saúde 2011).

Evidências científicas indicam que a detecção e a remoção precoce de pólipos pré-malignos reduzem a mortalidade do CCR, confirmadas por vários estudos que realizaram rastreamento de populações de risco geral através do teste de sangue oculto na fezes. Além disso, o rastreamento invasivo pelo uso da sigmoidoscopia flexível demonstrou resultados

promissores em estudos randomizados no Reino Unido (ATKIN et al. 2010) e nos Estados Unidos (SCHOEN et al. 2012), onde se observaram reduções significativas, tanto na incidência quanto na mortalidade do CCR. A melhora da sobrevida também foi observada com esta abordagem em grupos de alto risco geneticamente determinados (DE JONG et al. 2006). Por isso, àqueles indivíduos com alto risco poderia ser oferecida uma vigilância mais intensiva, com colonoscopia ou sigmoidoscopia flexível periodicamente em intervalos menores de tempo. A colonoscopia já é oferecida aos indivíduos com risco elevado devido à história pessoal ou familiar de CCR, assim como nas famílias com síndrome de Lynch e síndromes com polipose intestinal, recomenda-se a vigilância mais assídua (LEVIN et al. 2008). Portanto, estratificar a população geral em categorias de risco possibilitaria a individualização de estratégias de rastreamento e prevenção.

1.2 ETIOPATOGENESE

Devido à acessibilidade do intestino grosso para se detectar e rastrear lesões neoplásicas através de biópsias, o CCR é um dos cânceres mais estudados quanto à sua etiopatogênese.

A clássica descrição da sequência adenoma-carcinoma na tumorigênese do CCR por VOGELSTEIN et al. (1988) há mais de duas décadas, reflete na compreensão da patogênese molecular do CCR até os dias atuais. São três as principais vias moleculares: (1) a instabilidade cromossômica, responsável pela maioria dos casos esporádicos, onde há

rearranjos de vários segmentos cromossômicos (5q, 8p 17p e 18q), que contribuem para a aneuploidia tumoral (LENGAUER et al. 1998); (2) a instabilidade de microssatélites, presente em cerca de 15% dos casos esporádicos e na maior parte dos casos hereditários, sendo que nos primeiros deve-se à hipermetilação de ilhas CpG do promotor dos genes do sistema de reparo de mau-pareamento do DNA (amiúde o *MLH1*) (CUNNINGHAM et al. 1998); e (3) a via dos pólipos serrilhados, mais recente, que envolvem mecanismos moleculares diferentes da sequência clássica adenoma-carcinoma (NOFFSINGER 2009).

O CCR é uma neoplasia complexa, causada pela interação de múltiplos fatores, genéticos e ambientais. Entre os últimos estão a dieta rica em carne vermelha e gordura animal e pobre em fibras, o tabagismo e o excesso de álcool, a obesidade e o sedentarismo, e a doença inflamatória crônica intestinal (CUNNINGHAM et al. 2010). Além da idade, do sexo masculino e da história prévia de pólipos, a história familiar é considerada o principal fator de risco, sendo o risco relativo entre irmãos de duas a três vezes maior do que na população geral (CANNON-ALBRIGHT et al. 1988).

Um extenso estudo que analisou a taxa de concordância entre gêmeos monozigóticos e dizigóticos com diversos tipos de câncer definiu que a herdabilidade, isto é, a proporção da variância fenotípica explicada por fatores genéticos herdados, é de aproximadamente 35% (LICHTENSTEIN et al. 2000).

O CCR pode ser classificado, de maneira geral, em dois grupos: esporádico e familiar. Alguns estudos estimam que a proporção de casos

familiais seja de 10%, enquanto que em outros, representam pelo menos 25% de todos os casos (ST JOHN et al. 1993).

O conceito de CCR familiar reflete um extremo de um espectro de risco determinado pela contribuição de variantes genéticas de suscetibilidade. A maioria dos genes de suscetibilidade ao CCR foi identificada em famílias acometidas por síndromes hereditárias, as quais são causadas por mutações de alta penetrância. Essas síndromes representam cerca de 6% dos casos de CCR, e podem ser classificadas em síndromes com polipose ou sem polipose gastrointestinal (AALTONEN et al. 2007). Entre as principais síndromes com polipose destacam-se: a polipose adenomatosa familiar, causada por mutações no gene *APC*; a síndrome de Peutz-Jeghers, com mutações no gene *LKB1/STK11*; a polipose juvenil, associada ao gene *BMPR1A* e a síndrome de Cowden, ao gene *PTEN*. Entre as síndromes sem polipose, a mais prevalente é a síndrome de Lynch, representando cerca de 3% de todos os casos com CCR (AALTONEN et al. 1998), causada por mutações nos genes de reparo de mau-pareamento durante a replicação do DNA (*MLH1*, *MSH2*, *MSH6*, *PMS2* e *EPCAM*). O teste genético e aconselhamento de famílias com tais síndromes hereditárias têm auxiliado na identificação precoce de indivíduos com maior risco de desenvolver o CCR e outros tumores associados (*Evaluation of Genomic Applications in Practice and Prevention-EGAPP* 2009).

A maioria das mutações identificadas no CCR familiar são de alta penetrância, isto é, com grande chance de manifestar o câncer (ou cânceres) ao longo da vida. Entretanto, existem famílias com agrupamentos

de casos com CCR que não apresentam mutações nos genes associados às síndromes hereditárias. Isto levanta a hipótese de haver outras variantes ou mutações de baixa penetrância que tornem determinados indivíduos mais suscetíveis ao desenvolvimento do CCR. Estudos com irmãos que são concordantes ou discordantes para o CCR, assim como diversos estudos de associação, têm identificado regiões no genoma humano nas quais variantes do tipo polimorfismos de nucleotídeo único (*SNPs*, do inglês, *single nucleotide polymorphisms*) estão associados à suscetibilidade ao CCR (TOMLINSON et al. 2010).

1.3 POLIMORFISMOS DE NUCLEOTÍDEO ÚNICO E SUSCETIBILIDADE GENÉTICA

Os polimorfismos de nucleotídeo único ou *SNPs* são variações do genoma humano onde dois ou ocasionalmente três nucleotídeos alternativos são comuns na população. Na maioria dos casos, um *SNP* possui duas formas alternativas, denominadas alelos, por exemplo, A ou G numa certa posição do genoma.

Estima-se que haja cerca de dez milhões de *SNPs* no genoma humano, representando, juntamente com outros tipos de polimorfismos (como as variações do número de cópias) cerca de 90% da variação genética humana, incluindo a suscetibilidade a doenças. Quaisquer dois indivíduos são 99,5% idênticos em suas sequências de DNA e, a cada 1.000 pares de bases, há um *SNP* (SACHIDANANDAM et al. 2001).

O termo polimorfismo é utilizado por geneticistas humanos para significar várias coisas diferentes em tempos diferentes. Os geneticistas populacionais definem-no como a coexistência estável numa população de mais de um genótipo cujas frequências mais raras não possam manter-se simplesmente por mutações recorrentes. Já os geneticistas clínicos consideram polimorfismo como uma variante não patogênica, independentemente de sua frequência, enquanto que mutações são variantes patogênicas, sejam raras ou comuns. Por fim, os geneticistas moleculares designam como polimorfismo aquela variante cuja frequência na população está acima de um valor arbitrário, geralmente 1% (STRACHAN e READ 2011).

As variantes que foram deletérias durante a evolução (mutações que causam doenças precoces) são particularmente raras devido à seleção natural. Por sua vez, as variantes patogênicas que são deletérias em homozigose podem haver tornado-se neutras ou sofrido um equilíbrio de seleção por conferir uma vantagem em heterozigotos assintomáticos (como na anemia falciforme em relação ao patógeno da malária). Por isso, não se espera que alelos de *SNPs* frequentes na população tenham algum efeito fenotípico considerável, tanto porque a seleção natural se encarregaria de eliminá-la, se fosse prejudicial (seleção negativa), ou fixá-la, se benéfica (seleção positiva), quanto pela sua localização, já que a maioria dos *SNPs* não se encontra em sequências codificadoras ou reguladoras, podendo estar localizados a grandes distâncias de um gene cuja atividade seja afetada (STRACHAN e READ 2011).

Os fatores de suscetibilidade podem ser polimorfismos no DNA ou no RNA não-codificante (como os microRNAs), que produzem algum efeito menor na atividade promotora, no processamento ou na estabilidade do RNA mensageiro. Assim, tais variantes provavelmente afetam sutilmente a expressão ou função gênica, ao contrário das mutações deletérias (*frameshift*, *nonsense*, *splice-site* e *missense*) e frequentemente encontradas em doenças monogênicas (STRACHAN e READ 2011).

A suscetibilidade pode ser devido à combinação de diversas variantes ou independentemente. Uma ferramenta que pode desvendar a relação causal entre as variantes é a regressão logística, onde uma variante principal é selecionada, enquanto que os efeitos das outras variantes são condicionados nos efeitos da variante principal. Se ainda assim houver um efeito, a segunda variante contribui independentemente (STRACHAN e READ 2011).

Um problema a ser considerado é que os principais fatores de suscetibilidade podem ser diferentes em diferentes populações. Exceto as diferenças devido à interação dos fatores genéticos com os ambientais, as associações identificam segmentos cromossômicos ancestrais compartilhados, os quais podem diferir entre as populações (STRACHAN e READ 2011).

1.4 SUSCETIBILIDADE GENÉTICA AO CÂNCER COLORRETAL

Durante os últimos 15 anos, têm-se empregado um grande esforço para identificar genes adicionais de alta penetrância para o CCR, porém, poucos têm sido identificados atualmente. As análises estatísticas dos dados baseados nos padrões de agregação familiar do câncer, excluindo-se os genes de alto risco (citados anteriormente), sugerem que a maioria da suscetibilidade herdada ao câncer deve-se a um modelo poligênico de predisposição, no qual a herança conjunta de variantes genéticas, cada qual com um efeito individual modesto, podem causar uma ampla parcela do risco populacional. Estudos recentes têm identificado três classes principais de alelos de suscetibilidade ao câncer com diferentes níveis de risco e prevalência na população (FLETCHER e HOULSTON 2010):

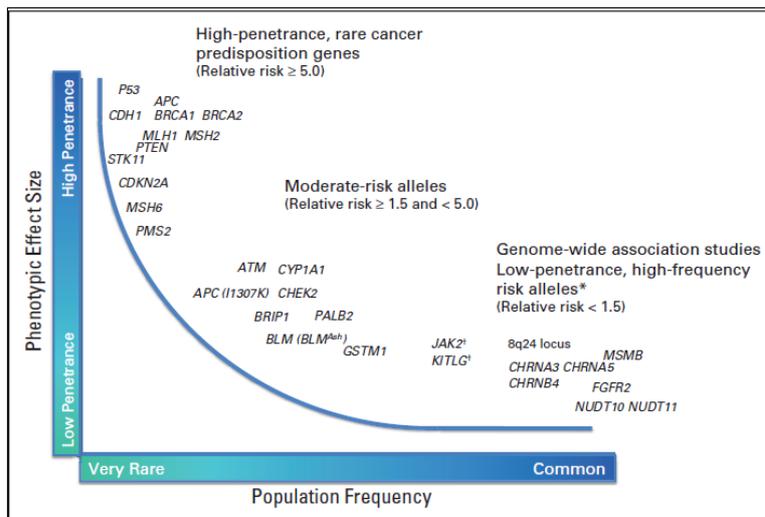
- Variantes raras de alta penetrância, responsáveis por síndromes hereditárias de predisposição ao câncer com risco relativo maior que cinco vezes ao da população geral;
- Variantes raras de penetrância moderada, com frequência do alelo mais raro menor que 2% e risco maior que o dobro da população geral;
- Variantes comuns de baixa penetrância, com frequência do alelo mais raro maior que 10% e risco menor que 1,5 vez da população geral.

Segundo a “hipótese da variante rara” uma proporção substancial da suscetibilidade genética ao câncer deve-se a adição dos efeitos dos alelos de risco com penetrância moderada, que incluem variantes subpolimórficas

e mutações (BODMER e BONILLA 2008). Estes alelos funcionam de forma dominante e independente, cada um conferindo um aumento moderado, mas prontamente identificável no risco, sendo, geralmente, específico de uma população, devendo-se a efeitos fundadores resultantes de deriva genética. Exemplo é a variante polimórfica I1307K no gene *APC*, presente em cerca de 6% dos judeus Ashkenazi, que confere cerca do dobro do risco do CCR e adenoma colorretal (LAKEN et al. 1997).

Há muito tempo especula-se que variantes polimórficas comuns contribuem com a suscetibilidade ao câncer, mas evidências dos alelos de baixa penetrância surgiram apenas recentemente.

Um grande número de evidências sugere que a contribuição genética no CCR deve-se a variantes comuns com efeitos individuais pequenos, invocando a hipótese 'doença comum-variante comum'. A teoria da genética populacional prediz que a distribuição dos efeitos alélicos que influenciam traços complexos é em forma de 'L', com um pequeno número de variantes responsáveis por um grande efeito no fenótipo da doença, enquanto que um grande número de variantes com pequeno efeito individualmente (Figura 1). Desse modo, a penetrância das variantes é inversamente proporcional à frequência alélica na população (BOST et al. 2001).



Fonte: STADLER et al. (2010)

Figura 1 - As três classes de alelos de suscetibilidade ao câncer.

Assim, as variantes raras de alta penetrância contribuem predominantemente no subgrupo de pacientes com segregação familiar da doença, no caso, do CCR. Este modelo implica que análises de associação baseadas no mapeamento de variantes polimórficas em todo o genoma, denominadas, em inglês, de *genome-wide association studies (GWAS)*, poderia ser uma estratégia robusta para identificar variantes comuns de baixo risco para o desenvolvimento do CCR (TOMLINSON et al. 2010).

Apesar das variantes causais não terem sido identificadas até o momento, interessante, poucos *SNPs* localizam-se em regiões codificantes, sendo que a maioria está distante de genes. Isto sugere que, provavelmente, tais variantes comuns modifiquem o risco de CCR ao influenciarem a expressão gênica, talvez de maneira sutil, ou pelo efeito em componentes de vias de sinalização mitigado por redundância funcional (TOMLINSON et al. 2010).

O primeiro locus de suscetibilidade ao CCR identificado foi 8q24, que havia sido previamente associado com o câncer de próstata e mama (AMUNDADOTTIR et al. 2006; EASTON et al. 2007). Sabe-se que nesta região não há qualquer gene ou transcritos codificantes, com a exceção de um pseudogene (*POU5F1P1*). Entretanto, aproximadamente a 300 kb telomérico desta região está o oncogene *c-MYC*. Os GWAS e estudos replicativos identificaram o *SNP* rs698327 como a variante mais promissora para validação funcional, e POMERANTZ et al. (2009) demonstraram que variantes nesta região são funcionalmente importantes provavelmente por regular a expressão de *c-MYC*.

Dos 20 *SNPs* identificados, pelo menos cinco deles estão próximos da família de genes da via de sinalização TGF- β e BMP, previamente implicada na patogênese tumoral, como *SMAD7*, *GREM1*, *BMP2*, *BMP4* e *RHNP2* (BRODERICK et al. 2007; JAEGER et al. 2008; HOULSTON et al. 2008). O gene *EIF3H* em 8q23.3 regula o crescimento e a viabilidade celular (TOMLINSON et al. 2008) e o gene *CRAC1* em 15q13.3 havia sido descrito em associação com a síndrome de polipose intestinal mista em indivíduos descendentes de judeus Ashkenazi (JAEGER et al. 2003). Pela primeira vez, um *SNP* (rs 5934683) no cromossomo X (Xp22.2) foi associado à predisposição genética ao CCR, um câncer não relacionado ao sexo (DUNLOP et al. 2012a).

Elucidar os efeitos funcionais das variantes comuns na expressão de componentes da via TGF- β será uma tarefa desafiadora para aumentar o

conhecimento da etiopatogênese do CCR, e, conseqüentemente, levar a novas perspectivas terapêuticas.

1.5 ESTUDOS DE ASSOCIAÇÃO DE LARGA ESCALA GENÔMICA

Pesquisar associações populacionais é uma opção atrativa para identificar genes de suscetibilidade a doenças. Os estudos de associação são mais fáceis de serem conduzidos do que os estudos de ligação, pois não necessitam de múltiplos casos familiares segregando o fenótipo. Entretanto, dependem do desequilíbrio de ligação (DL) - a associação não-randômica entre alelos em diferentes *loci* - com um fator de suscetibilidade, o qual só pode ser identificado por marcadores localizados no mesmo bloco do haplótipo - o conjunto de alelos num *loci* ligados em um único cromossomo - próximo do fator (VISSCHER et al. 2012).

Os SNPs são os marcadores de escolha para estudar o DL por três razões: (1) são suficientemente abundantes que permitem verificar segmentos cromossômicos muito curtos; (2) em comparação com os microssatélites, eles possuem uma taxa menor de mutação; (3) os SNPs são facilmente genotipados em larga escala do genoma (STRACHAN e READ 2011).

A estrutura do DL humano foi investigada pelo projeto *HapMap* (*Haplotype Map of the Human Genome*), e o primeiro resultado foi uma lista de mais de 3 milhões de *SNPs* que capturaram a maioria da variação

genômica comum em algumas populações (*International HapMap Consortium 2005*).

Os *GWAS* são embasados no princípio do DL em nível populacional, que, normalmente, é o resultado de um determinado haplótipo ancestral comum numa população. Geralmente, os *loci* que estão fisicamente próximos exibem um DL mais forte do que os que estão distantes em um cromossomo. A distância genômica à qual o DL decai determina quantos marcadores genéticos são necessários para ‘etiquetar’ um bloco haplotípico, sendo o número de tais marcadores muito menor do que o total de variantes segregando na população. Por exemplo, a seleção de cerca de 500 mil *SNPs* comuns no genoma humano é suficiente para ‘etiquetar’ as variantes comuns nas populações não-Africanas, mesmo que o total de *SNPs* seja maior que 10 milhões (*International HapMap Consortium 2005*). Estes *SNPs* são chamados de ‘etiquetas’ (do inglês, *tagSNPs*).

Com a utilização de densas plataformas de genotipagem de alto desempenho, os *SNP arrays*, é possível genotipar vários marcadores em um único ensaio. O uso de biobancos populacionais de amostras de casos e controles têm facilitado aos *GWAS* a obtenção de resultados mais rápidos e custo-efetivos com reprodutibilidade (TENESA e DUNLOP 2009).

Embora os *GWAS* não sejam influenciados pelo conhecimento prévio biológico nem da localização genômica dos *SNPs*, são influenciados pelo DL entre os *SNPs* genotipados e as variantes causais não-genotipadas. A força da associação estatística entre os alelos nos dois *loci* no genoma depende, principalmente, das suas frequências alélicas. Assim, uma variante rara

(frequência do alelo *minor* – FAm - menor que 0,01) estará em baixo DL com uma variante comum vizinha, mesmo que estejam em um mesmo intervalo de recombinação. Porém, a maioria dos *SNPs* selecionados nos *SNP arrays* são comuns (FAm maior que 0,05) e por isso, os *GWAS* têm poder para detectar associação de variantes que são relativamente comuns na população (VISSCHER et al. 2012). Por outro lado, sugere-se que a associação observada entre um *SNP* comum e um traço complexo possa resultar do DL do *SNP* com variantes raras no mesmo locus. Como os alelos comuns e as variantes raras causais estão correlacionados em um baixo DL, a hipótese de uma associação “sintética” implica que a magnitude do efeito das variantes causais é muito maior (de 50 a 100 vezes) do que a dos *SNPs* comuns genotipados pelos *GWAS*. Por exemplo, se um *SNP* explica 0,1% da variância fenotípica na população, a variante causal explicaria de 5 a 10% (DICKSON et al. 2010).

Os primeiros resultados dos *GWAS* foram publicados entre 2005 e 2006, mas o grande marco foi o estudo do *Welcome Trust Case Control Consortium* publicado na revista *Nature* em 2007 (*Wellcome Trust Case Control Consortium 2007*). Nesses cinco anos, mais de 2.000 *loci* foram significativamente associados com um ou mais traços complexos, como doenças psiquiátricas, doenças inflamatórias intestinais, diabetes, e até mesmo estatura (*National Human Genome Research Institute 2012*). Por exemplo, nas populações europeias, estima-se que a maioria das variantes expliquem cerca de 10% da agregação familiar no diabetes tipo 2 (VISSCHER et al. 2012).

1.5.1 Estudos de associação de larga escala genômica e câncer colorretal

Embora a suscetibilidade genética ao CCR seja conhecida, o mapeamento dos alelos de predisposição tem sido desafiador. Avanços a partir da metade de 2007, por meio da abordagem dos *GWAS*, têm permitido a identificação de variantes comuns envolvidas na etiologia do CCR.

A maioria dos estudos, para identificar alelos de baixa penetrância para suscetibilidade ao CCR, foram embasados numa abordagem de genes candidatos, cujo papel na patogênese do CCR era supostamente conhecido. Entretanto, sem o real entendimento da biologia da predisposição, a escolha dos genes era problemática. Assim, até o advento dos *GWAS*, poucos ou quase nenhum estudo de associação baseado nesta abordagem - a maioria envolveu genes relacionados ao metabolismo de carcinógenos da dieta e reparo do DNA - foi capaz de identificar alelos de suscetibilidade inequivocamente associados ao risco de CCR (TOMLINSON et al. 2010).

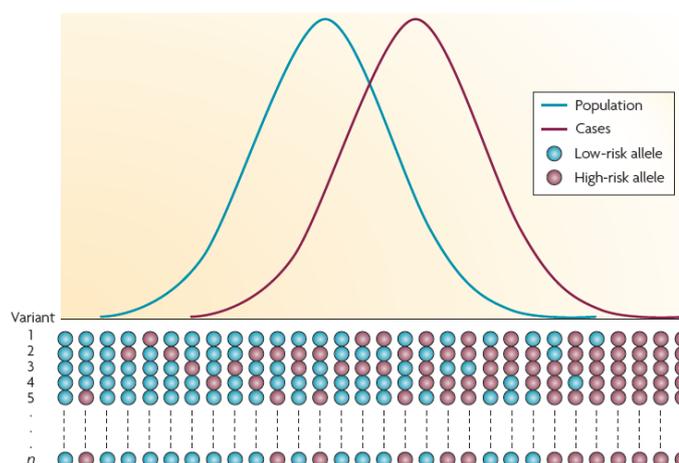
Até o momento, sete *GWAS* e três metanálises identificaram cerca de 20 *loci* independentes associados com a suscetibilidade ao CCR (ZANKE et al. 2007; TOMLINSON et al. 2007; HAIMAN et al. 2007; BRODERICK et al. 2007; JAEGER et al. 2008; TENESA et al. 2008; TOMLINSON et al. 2008; HOULSTON et al. 2008; HOULSTON et al. 2010; DUNLOP et al. 2012a; PETERS et al. 2012 (Tabela 1).

Tabela 1 - Dados dos GWAS em CCR com os SNPs identificados até o momento.

SNP ID	Locus/ Cromossomo	Localização/Genes mais próximos	População	Frequência do alelo de risco em controles	OR alélico (IC95%)	Referência
rs6983267*	8q24.21	intergênica/MYC	Europeia (RU)	51%	1.21 (1,15 – 1,27)	Tomlinson et al. 2007
rs4939827*	18q21.1	SMAD7 (intron 3)	Europeia (RU)	47-52%	1.18 (1,12 – 1,23)	Broderick et al. 2007
rs4779584*	15q13.3	intergênica/SCG5, GREM1	Europeia (RU)	18%	1.26 (1,19 – 1,34)	Jaeger et al. 2008
rs16892766*	8q23.3	intergênica/EIF3H	Europeia (RU)	7%	1.25 (1,19 – 1,32)	Tomlinson et al. 2008
rs10795668*	10p14	Intergênica/BC031880	Europeia (RU)	67%	1.12 (1,10 – 1,16)	Tomlinson et al. 2008
rs4444235*	14q22.2	Intergênica/BMP4	Europeia (RU)	46%	1.11 (1,08 – 1,15)	Houlston et al., 2008
rs9929218*	16q22.1	CDH1 (intron)	Europeia (RU)	71%	1.10 (1,06 – 1,12)	Houlston et al., 2008
rs10411210*	19q13.1	RHPN2 (intron)	Europeia (RU)	10%	1.15 (1,10 – 1,20)	Houlston et al., 2008
rs961253*	20p12.3	Intergênica/BMP2	Europeia (RU)	36%	1.12 (1,08 – 1,16)	Houlston et al., 2008
rs3802842*	11q23.1	Intergênica/FLJ45803	RU/Canadá/ Israel/Japão	29%	1.12 (1,07 – 1,17)	Tenesa et al. 2008
rs6691170	1q41	254Kb upstream DUSP10	Europeia (RU, Finlândia)	34-37%	1,06 (1,03-1,09)	Houlston et al. 2010
rs6687758	1q41	125Kb upstream DUSP10	Europeia (RU, Finlândia)	19-26%	1,09 (1,06-1,12)	Houlston et al. 2010
rs10936599	3q26.2	MYNN	Europeia (RU, Finlândia)	24-28%	0,93 (0,91-0,96)	Houlston et al. 2010
rs7136702	12q13.13	telomérico ao LARP4 e proximal ao DIP2	Europeia (RU, Finlândia)	29-37%	1,06 (1,04-1,08)	Houlston et al. 2010
rs11169552	1213.3	telomérico ao DIP2B e proximal ao ATF1	Europeia (RU, Finlândia)	26-41%	0,92 (0,90-0,95)	Houlston et al. 2010
rs4925386	20q13.33	LAMA5	Europeia (RU, Finlândia)	30-33%	0,93 (0,91-0,95)	Houlston et al. 2010
rs1957636	14q22.2	BMP4	Europeia	n.i.	1,08 (1,06 – 1,11)	Tomlinson et al. 2011
rs4813802	20p12.3	BMP2	Europeia	n.i.	1,09 (1,06 – 1,12)	Tomlinson et al. 2011
rs7315438	12	MED13L	EUA/Canadá	42%	0,90 (0,87 – 0,94)	Peters et al. 2012
rs275454	5	LOC729434/POLS	EUA/Canadá	37%	1,08 (1,04 – 1,13)	Peters et al. 2012
rs2853668	5	TERT-CLPTM1L	EUA/Canadá	25%	0,91 (0,87 – 0,96)	Peters et al. 2012
rs2373859	2	SLC8A1	EUA/Canadá	35%	0,93 (0,9 – 0,97)	Peters et al. 2012
rs1525461	7	LOC643308/TPK1	EUA/Canadá	19%	1,09 (1,04 – 1,14)	Peters et al. 2012
rs1321311	6p21	CDKN1A	Europeia/ Japonesa	n.i.	1,10 (1,07-1,13)	Dunlop et al. 2012a
rs3824999	11q13.4	POLD3 (intron 9)	Europeia/ Japonesa	n.i.	1,08 (1,05 – 1,10)	Dunlop et al. 2012a
rs5934683	Xp22.2	entre GPR143 e SHROOM2	Europeia/ Japonesa	n.i.	1,07 (1,04 – 1,10)	Dunlop et al. 2012a

* SNPs avaliados neste estudo; OR = odds ratio; n.i. = não informado; RU = Reino Unido

Os riscos associados a cada dos *SNPs* localizados nesses *loci* são modestos (*odds ratio* 1,1 a 1,3), havendo pouca evidência de efeitos interativos entre eles ou epistase. Os riscos das variantes em homozigose conferem o dobro do risco das variantes em heterozigose, sendo que os alelos de risco seguem uma curva de distribuição normal, tanto em casos quanto em controles, com um desvio em direção a um maior número de alelos de risco entre os casos, o que é consistente com um modelo poligênico de predisposição a doenças comuns (Figura 2) (FLETCHER e HOULSTON 2010).



Fonte: FLETCHER e HOULSTON (2010).

Figura 2 - A distribuição dos alelos de risco em casos e controles.

O número de variantes comuns que contribuem com mais de 1% do risco herdado é muito baixo, sendo que é muito improvável que haja outros *SNPs* com efeitos similares (maior do que 1,2) para alelos com frequências maiores do que 20% em populações europeias. Com efeito, os *GWAS* identificaram em média 80% dos *SNPs* comuns nessa população, mas

somente 12% dos *SNPs* com frequência do alelo *minor* (F_{Am}) entre 5 e 10% (TOMLINSON et al. 2010).

Entretanto, as variantes com esse perfil, se tomadas coletivamente, podem conferir riscos substanciais devido à sua multiplicidade, e no caso do CCR, até o momento, explicam cerca de 10% da herdabilidade (DUNLOP et al. 2012a). Num modelo construído com base em dados do *GWAS* escocês, cerca de 170 variantes independentes comuns explicariam toda a variância genética do CCR (TENESA e DUNLOP 2009). Portanto, a maior parte da suscetibilidade genética ao CCR ainda precisa ser definida, a chamada herdabilidade faltante (do inglês, *missing heritability*). Existem outras possíveis causas dessa parcela não identificada: (1) o efeito das variantes raras; (2) a falha em identificar as variantes causais e (3) a heterogeneidade alélica (EHRET et al. 2012).

As estratégias dos *GWAS* de identificar alelos comuns que conferem riscos modestos não são ideais para identificar variantes raras (F_{Am} abaixo de 1%) com efeitos potencialmente maiores, assim como para capturar variantes de número de cópias e outras variantes estruturais, como inserções, rearranjos complexos ou expansões de repetições de microssatélites, as quais podem alterar o risco de CCR. Na medida em que se empreendam esforços para expandir a escala das metanálises dos *GWAS*, tanto em termos de tamanho da amostra quanto em cobertura de *SNPs*, assim como para aumentar o número de *SNPs* considerados para replicação em larga escala, será factível a descoberta de novas variantes. É possível que uma abordagem de múltiplos *loci* baseada em haplótipos de

marcadores identifique alelos raros. Além disso, o uso do sequenciamento de exoma pode fornecer uma estratégia mais eficaz para encontrar tais variantes (FLETCHER e HOULSTON 2010).

1.6 IMPLICAÇÕES FUTURAS

Apesar da magnitude do efeito sobre o risco atribuído a cada um desses *SNPs* ser pequena, coletivamente, contribuem para uma grande parcela do risco, devido às suas altas frequências alélicas na população. Assim, naqueles subgrupos populacionais portadores de múltiplos alelos de risco, o alto risco absoluto pode exceder os limiares de uma intervenção clínica. Por isso, a genotipagem de variantes comuns em uma população oferece a possibilidade de estratificar o risco conferido por múltiplos genótipos dentro do risco populacional geral, de acordo com a frequência desses genótipos na população. Assim, a individualização das estratégias preventivas, como uma vigilância mais intensiva poderia ser oferecida àqueles com maior risco, enquanto que aqueles com risco geral seriam rastreados de acordo com os protocolos regulares, assim como proposto no câncer de mama (TENESA e DUNLOP 2009).

DUNLOP et al. (2012b) realizaram um estudo com mais de 40.000 indivíduos (cerca de 25.000 casos e 20.000 controles) de sete populações predominantemente europeias com o objetivo de avaliar a exequibilidade da predição de risco do CCR. Foram desenvolvidos e validados modelos de predição de risco, combinando os dados de genotipagem de dez *SNPs* de

suscetibilidade, idade, gênero e história familiar. Em média, os casos tiveram um maior número de alelos de risco que os controles, e cada alelo aumentou 9% do risco, em média. O efeito combinado da genotipagem, história familiar, idade e gênero foi eficaz para identificar um subgrupo populacional com risco absoluto maior do que 5% em 10 anos.

Segundo um modelo de predição de risco, seria necessário identificar cerca de 100 *SNPs* para predizer com 80% de acurácia o risco genético do CCR. Utilizando apenas os dez *SNPs* pesquisados, a acurácia da predição é de apenas 26% (TENESA e DUNLOP 2009).

Em conclusão, a predição do risco baseando-se somente nos genótipos ainda não é viável, porém, ao incorporar informações de outros fatores de risco, como idade, gênero e história familiar, melhorar-se-ia o desempenho discriminatório do modelo de predição. A aplicação de tais modelos poderia auxiliar na estratificação populacional em grupos de alto risco, que mereceriam um rastreamento mais intensivo (com colonoscopias mais frequentes, por exemplo). Desse modo, portanto, seria possível identificar subgrupos populacionais com um risco de CCR suficientemente alto e clinicamente relevante, capaz de modificar as estratégias de saúde pública (FLETCHER e HOULSTON 2010).

O presente estudo faz parte de um projeto colaborativo multicêntrico internacional intitulado *Genetic Study of Common Hereditary Bowel Cancers in Hispania and the Americas* (CEP 1169/08; CONEP 15245), que tem por objetivo principal caracterizar a população latino-americana quanto à sua ancestralidade e suas variantes genéticas de predisposição ao CCR.

2 OBJETIVOS

O **objetivo geral** do presente estudo foi identificar os dez *SNPs* associados ao risco de CCR que foram descritos previamente em populações europeias em indivíduos da população brasileira.

Os **objetivos específicos** foram:

1. calcular as frequências alélicas e genotípicas dos dez *SNPs* em casos e controles;
2. analisar a associação dos genótipos e alelos dos dez *SNPs* com o risco de CCR;
3. calcular a magnitude do efeito sobre o risco de CCR; e
4. correlacionar os genótipos dos dez *SNPs* com as características clínico-patológicas e a história familiar.

3 MATERIAL E MÉTODOS

3.1 CASUÍSTICA

Este é um estudo retrospectivo de associação genética caso-controle, cuja amostra inicial compreendeu 787 casos e 750 controles, recrutados dos departamentos de Cirurgia Pélvica, Oncologia Clínica e Medicina Comunitária do Hospital do Câncer A.C. Camargo (HACC). Todos os pacientes e controles autorizaram a presente pesquisa mediante a assinatura do termo de consentimento livre e esclarecido previamente aprovado pelo Comitê de Ética em Pesquisa do HACC sob o número 1231/09 (Anexos 1 a 3).

Foram adotados os seguintes critérios de **inclusão**:

- Casos: pacientes com diagnóstico de adenocarcinoma de cólon e/ou reto antes dos 75 anos de idade ou com adenoma colorretal avançado (com histologia vilosa ou maior do que 1 cm ou com displasia severa) diagnosticado antes dos 60 anos.
- Controles: indivíduos sem CCR ou adenoma colorretal que não apresentassem parentes de primeiro ou segundo grau com CCR.

Os controles não foram pareados com os casos em relação à condição sócio-econômica e ancestralidade ou etnia auto-referida.

Os critérios de **exclusão** foram: (1) síndromes hereditárias de predisposição ao CCR; (2) imuno-histoquímica com ausência de proteínas

dos genes de reparo de mau-pareamento do DNA; (3) presença de mutações germinativas de alta penetrância em genes de suscetibilidade ao CCR; (4) tumores do apêndice; (5) doença inflamatória crônica intestinal prévia.

Inicialmente foram coletadas amostras de 787 casos, das quais, 418 foram provenientes do Banco de Macromoléculas do HACC, sendo 86 amostras compartilhadas com outro projeto de pesquisa (CEP nº 1542/11). Foram excluídas 60 amostras, 36 por preencherem os seguintes critérios de exclusão: onze eram maiores que 75 anos; sete apresentavam imunohistoquímica com ausência de expressão em pelo menos um gene de reparo de mau-pareamento do DNA; seis eram indivíduos com poliposes adenomatosas familiares; cinco preenchiam critérios clínicos de HNPCC; três tinham tumores neuroendócrinos; dois com tumores do apêndice; um com critério clínico da síndrome de Li-Fraumeni; e um com doença inflamatória crônica intestinal prévia. Seis amostras foram excluídas por falta de qualidade do DNA, quatro foram coletas repetidas e quatorze não obtiveram genótipos bem agrupados em mais da metade dos *SNPs* ou cuja taxa de genotipagem foi menor do que 95%.

Do total de 750 amostras controles coletadas, nove foram excluídas por não obterem genótipos bem agrupados em mais da metade dos *SNPs* ou cuja taxa de genotipagem foi menor do que 95% e uma amostra repetida.

Foram pesquisadas as seguintes informações clínico-patológicas dos casos selecionados: critérios revisados de Bethesda; idade ao diagnóstico; sexo; estadiamento TNM clínico inicial (pré-neoadjudância) ou patológico

(sem neoadjuvância); topografia da lesão (cólon proximal, distal e reto); grau de diferenciação (pouco, moderadamente e bem diferenciado); tipo histológico (adenocarcinoma tubular/túbulo-viloso; adenocarcinoma mucinoso/diferenciação mucinosa; adenoma avançado); status atual da doença (vivo com doença, vivo sem doença, óbito por câncer, óbito sem câncer) e história familiar de CCR em parentes de até 3º grau (ANEXO 4).

3.2 METODOLOGIA

3.2.1 Coleta e processamento das amostras

Após a assinatura do termo de consentimento foi realizada a coleta de sangue periférico em dois tubos do tipo *vacutainer* com EDTA. Todas as amostras foram cadastradas no Laboratório de Macromoléculas do HACC (sistema *Biobank*) onde foram feitas a extração e purificação do DNA genômico com a utilização do *PureGene Blood Core kit* e *QiAmp DNA Blood Mini QIACUBE kit*, segundo as instruções do fabricante (QIAGEN, Hilden, Alemanha). A qualidade e concentração do DNA extraído foram verificadas em espectrofotômetro *NanoDrop ND-1000* (Thermo Scientific, Wilmington, EUA) cuja pureza foi considerada adequada se $OD_{260}/OD_{280} = 1,8-2,0$. As alíquotas foram diluídas a uma concentração de 5 nanogramas por microlitro e armazenadas em freezer a -20°C .

3.2.2 Amplificação do DNA

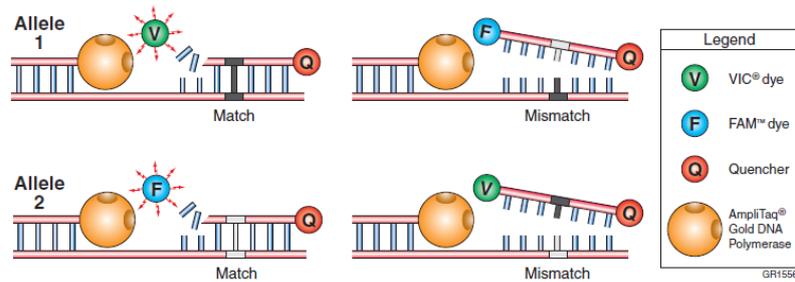
O DNA genômico de cada amostra foi amplificado por reação em cadeia de polimerase em tempo real (qPCR) na plataforma *7900HT Fast Real-Time PCR System* (Applied Biosystems, Foster City, EUA) do Laboratório de Genômica e Biologia Molecular do Centro de Pesquisas do HACC. As reações foram feitas em duplicata com um volume total de 9 microlitros contendo: 2,5 microlitros do *TaqMan® Genotyping Master Mix®*(Applied Biosystems, Foster City, EUA); 0,25 microlitro do *SNP Genotyping Assay®*(Applied Biosystems, Foster City, EUA), 4 microlitros da amostra *template* (DNA) e 2,25 microlitros de água Milli-Q autoclavada. Em cada ensaio foram adicionadas pelo menos duas amostras controles sem *template* e controles internos positivos para cada genótipo. As condições das reações de PCR foram: 95°C por 10 minutos, 92°C por 15 segundos e 60°C por 1 minuto por 40 ciclos. As reações foram pipetadas em placas de 384 poços (Applied Biosystems, Foster City, EUA) pelo pipetador automático QIAGILITY® (QIAGEN, Hilden, Alemanha) do laboratório NEOGENE do Centro de Pesquisas do HACC.

3.2.3 Genotipagem dos SNPs

Os seguintes *SNPs* previamente desenhados e validados foram genotipados: rs4939827, rs10411210, rs4444235, rs9929218, rs4779584, rs961253, rs10795668, rs6983267, rs3802842 e rs16892766. Por razões gráficas, os *SNPs* foram renumerados, respectivamente: 06, 09, 16, 21, 26, 45, 59, 71, 82 e 83.

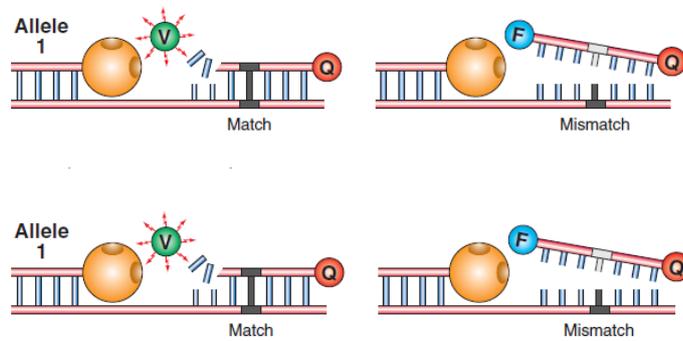
Os ensaios de genotipagem utilizaram o método *TaqMan*® *SNP Genotyping Assays* (Applied Biosystems, Foster City, EUA) e as análises de discriminação alélica foram feitas na plataforma *7900HT Fast Real-Time PCR System* pelo programa *Sequence Detection Systems Automation Controller Software v2.3* (Applied Biosystems, Foster City, EUA) mediante a técnica de *auto-calling*.

Os ensaios de discriminação alélica são testes multiplexados, isto é, utilizam mais de um par de iniciadores e sondas por reação, de ponto final (os dados são coletados ao final da reação de PCR), que detectam variantes de um único nucleotídeo na sequência do DNA pelo método fluorogênico 5' nuclease (LEE et al. 1993). A presença de dois pares de iniciadores/sondas em cada reação permite a genotipagem das duas variantes possíveis no sítio do *SNP* numa sequência-alvo da amostra. Para cada amostra, um par único de corantes fluorescentes é usado, VIC® e FAM™ (Applied Biosystems, EUA), onde o corante VIC® acopla-se perfeitamente no alelo selvagem ou *major* (alelo 1) e o FAM™ no alelo mutante ou *minor* (alelo 2). Os ensaios de discriminação alélica medem a variação na fluorescência dos corantes associados às sondas, determinando, assim, os três genótipos possíveis (heterozigotos, homozigotos raros e homozigotos comuns), conforme as figuras a seguir:



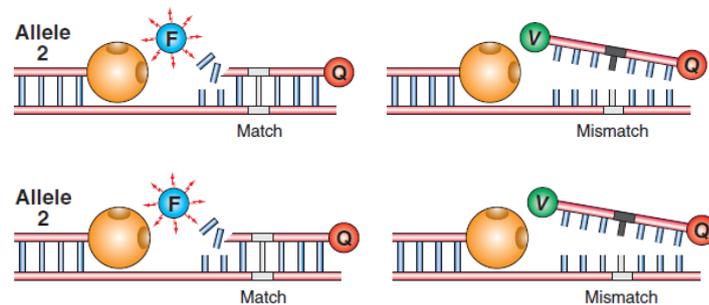
Fonte: Applied Biosystems

Figura 3 - Representação esquemática da genotipagem de um indivíduo heterozigoto (ambas as sondas emitem fluorescência)



Fonte: Applied Biosystems

Figura 4 - Representação esquemática da genotipagem de um indivíduo homozigoto para o alelo 1 (somente a sonda VIC® emite fluorescência)



Fonte: Applied Biosystems

Figura 5 - Representação esquemática da genotipagem de um indivíduo homozigoto para o alelo 2 (somente a sonda FAM emite fluorescência).

Como controle de qualidade da genotipagem, foram excluídos os genótipos com os seguintes critérios: (1) taxa global de genotipagem (*overall call rate*) menor do que 95%; (2) duplicatas discordantes; (3) desequilíbrio de Hardy-Weinberg ($p < 10^{-6}$ nos casos e $p < 10^{-4}$ nos controles); e (4) agrupamentos (*clusters*) de genótipos insatisfatórios na inspeção dos eixos X:Y.

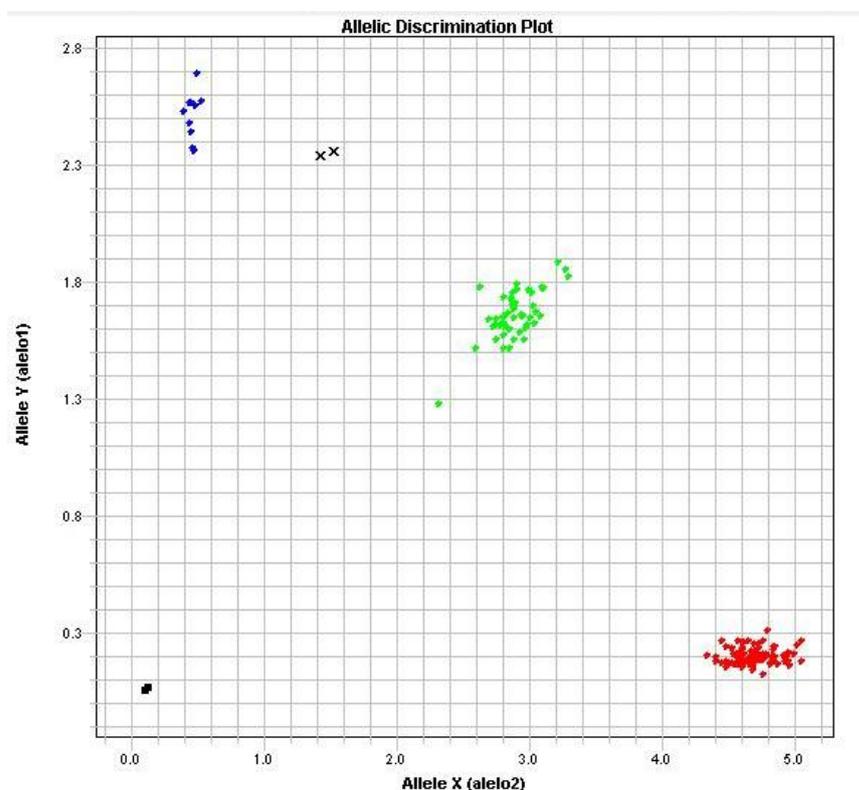


Figura 6 - Exemplo de *clusters* satisfatórios das amostras representando os três genótipos (em azul, homozigotos para o alelo 1; em verde, heterozigotos; em vermelho, homozigotos para o alelo 2). As cruces pretas representam uma amostra em duplicata cuja taxa de genotipagem foi abaixo de 95% e os quadrados pretos representam os controles negativos.

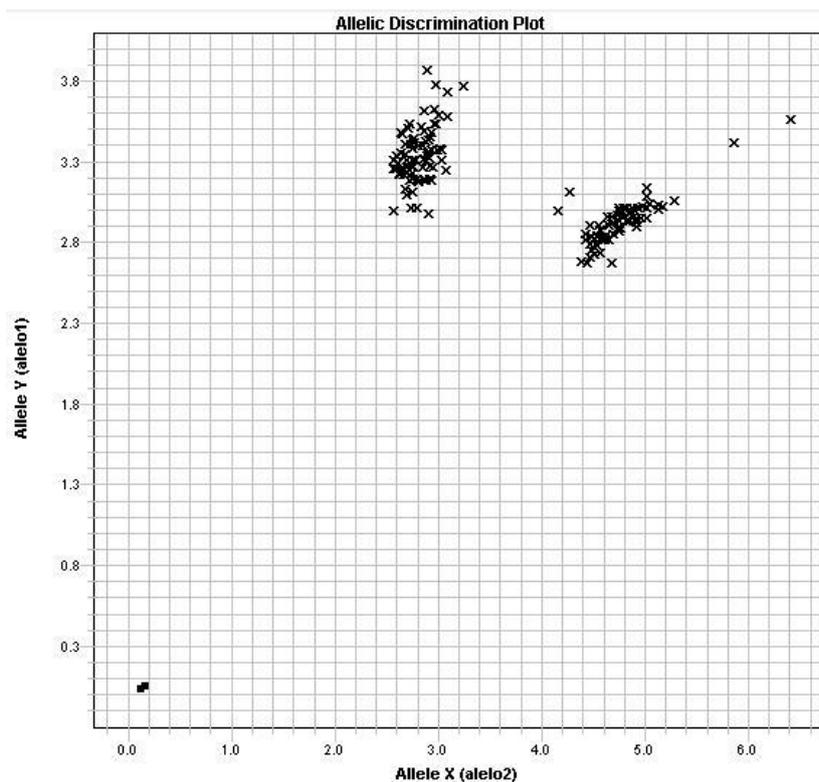


Figura 7 - Exemplo de *clusters* insatisfatórios. Não houve discriminação dos três genótipos.

3.3 ANÁLISES ESTATÍSTICAS

As frequências alélica e genotípica de cada *SNP* foram calculadas com a utilização do programa DeFinetti (<http://ihg2.helmholtz-muenchen.de/cgi-bin/hw/hwa2.pl>), assim como os desvios das frequências genotípicas nos casos e controles esperadas pelo equilíbrio de Hardy-Weinberg foram calculados pelo teste de qui-quadrado com um grau de liberdade ou pelo teste exato de Fisher, se a contagem esperada de células fosse menor que cinco.

As análises de associação entre os genótipos encontrados nos casos e controles para cada *SNP* foram realizadas com vários tipos de modelos genéticos descritos a seguir, com a utilização do programa *SNP & Variation Suite Version 7.6.10 (Golden Helix 2012)*:

- a. **Teste alélico básico:** os alelos (*major* ou referência e *minor* ou alternativo) são testados individualmente;
- b. **Teste genotípico:** os genótipos (homozigotos comuns, homozigotos raros e heterozigotos) são testados independentemente da ordem ou contagem de alelos;
- c. **Modelo aditivo:** a associação é testada dependendo da contagem do alelo *minor* ou alternativo, isto é, quando há dois alelos *minor* ou alternativos (homozigoto raro) ao invés de nenhum (homozigoto comum), duplica-se o efeito quando se compara com a presença de apenas um alelo *minor* ou alternativo (heterozigoto);
- d. **Modelo dominante:** a associação é testada pela presença de pelo menos um alelo *minor* ou alternativo (heterozigoto ou homozigoto raro) comparada com a ausência desse alelo (homozigoto comum);
- e. **Modelo recessivo:** associação é testada pela presença de dois alelos *minor* ou alternativos (homozigoto raro) comparada com a presença de pelo menos um alelo *major* ou referência (heterozigoto ou homozigoto comum)

Para as análises de associação genética citadas acima foram utilizados os seguintes testes estatísticos:

- a. **Teste de tendência Cochran-Armitage:** utilizado especificamente para o modelo aditivo, onde os genótipos dos casos são comparados com os dos controles seguindo uma tendência, a qual depende do número de alelos *minor* ou alternativos. A vantagem deste teste é que não é necessário assumir o equilíbrio de Hardy-Weinberg.
- b. **Teste de qui-quadrado de Pearson:** utilizado para qualquer modelo genético, exceto o aditivo, em tabelas de contingência descritas a seguir:
- Alélico básico: (casos versus controles) / (alelo *major* versus alelo *minor*) em tabelas 2x2;
 - Genotípico: (casos versus controles) / (homozigotos raros versus homozigotos comuns versus heterozigotos) em tabelas 2x3;
 - Dominante: (casos versus controles) / ([homozigotos raros ou heterozigotos] versus homozigotos comuns) em tabelas 2x2;
 - Recessivo: (casos versus controles) / (homozigotos raros versus [heterozigotos ou homozigotos comuns]) em tabelas 2x2
- c. **Teste exato de Fisher:** utilizado para qualquer modelo genético, exceto o aditivo, seguindo as mesmas tabelas de contingência que o teste qui-quadrado, porém evitando a aproximação do qui-quadrado ao analisar a probabilidade exata sob a hipótese nula de haver uma tabela de contingência pelo menos tão extrema quanto à observada,

assumindo que haja uma probabilidade igual de qualquer permutação da variável dependente.

d. **Razão de chances (*odds ratio*) com intervalos de confiança de 95%:** utilizada para qualquer modelo genético, exceto o genotípico como a seguir:

- Alélico básico: razão de chances do alelo *minor* ou alternativo aumentar o efeito e razão de chances do alelo *major* ou referência aumentar o efeito;
- Dominante: razão de chances “normal” ([homozigotos raros ou heterozigotos] versus homozigotos comuns) e razão de chances invertida (homozigotos comuns versus [homozigotos raros ou heterozigotos]);
- Recessiva: razão de chances “normal” (homozigotos raros versus [heterozigotos ou homozigotos comuns]) e razão de chances invertida ([heterozigotos ou homozigotos comuns] versus homozigotos raros)
- Aditivo: razão de chances dos heterozigotos versus homozigotos comuns e razão de chances dos homozigotos raros versus os heterozigotos. Este teste serve como validação do próprio modelo, ao indicar a intensidade do efeito sobre a associação. Assim, se as duas razões de chances são semelhantes, então o modelo aditivo é considerado válido. No entanto, se forem muito diferentes, então outro modelo seria mais adequado. Por exemplo, uma razão de chances pouco

significativa para homozigotos raros versus heterozigotos e muito significativa para heterozigotos versus homozigotos comuns indica que o melhor modelo para descrever o efeito seja o dominante.

- e. **Regressão logística:** utilizada em qualquer modelo genético, exceto o genotípico.

As análises das correlações dos genótipos com as variáveis clínico-patológicas e a história familiar foram feitas com o modelo de regressão logística binomial e análise univariada incondicional com o cálculo do *odds ratio* e intervalo de confiança, utilizando o programa *SNP Variation Suite version 7.6.10 (Golden Helix 2012)*. Foram selecionadas como variáveis dependentes os dados clínico-patológicos e a história familiar, as quais foram convertidas em variáveis binárias e, como variáveis independentes, cada um dos três genótipos de cada *SNP*. O nível de significância α foi de 0,05 (bicaudado). Os métodos de correção empregados para múltiplos testes para controlar o erro tipo I ou falsos positivos foram o de Bonferroni e a taxa de falsa descoberta (*false discovery rate*), que é a proporção de falsos positivos entre todos os positivos sob as hipóteses nula e alternativa (BALDING 2006).

4 RESULTADOS

4.1 CARACTERÍSTICAS DOS CASOS E CONTROLES

Dos 727 casos incluídos no estudo, 51% foram do sexo masculino, com média da idade ao diagnóstico de $56,9 \pm 10,1$ anos; 30% preencheram os critérios de Bethesda; 3% dos tumores foram adenomas avançados; a localização mais comum do CCR foi o reto e em torno de 10% houve um segundo tumor primário extra-cólon; o estadiamento III foi o mais frequente ao diagnóstico, com a maior parte dos tumores invadindo estruturas pericólicas/perirretais, porém, sem metástases à distância; predominaram os tumores moderadamente diferenciados e com tipo histológico tubular; porém em quase metade dos casos os adenocarcinomas não foram especificados. A maior parte dos pacientes estava viva sem doença até o final da coleta dos dados (agosto de 2012) e cerca de 30% dos casos não apresentava história familiar de CCR até parentes de 3^o grau, apesar de que cerca de 20% não souberam referir com exatidão (Tabela 2).

Dos 740 controles incluídos no estudo, 52% foram do sexo feminino, com média de idade de $51,9 \pm 12,3$ anos. Não houve diferença estatisticamente significativa entre os casos e controles quanto ao sexo ($p = 0,193$).

Tabela 2 - Características dos casos do Hospital A.C.Camargo entre 1990-2012.

Características	Categoria	N	%
Gênero	masculino	372	51,2
	feminino	355	48,8
Idade	≥ 50 anos	565	77,7
	< 50 anos	162	22,3
Critérios revisados de Bethesda	Sim	210	29,8
	Não	494	70,2
	n.a.	23	
Tumor	CCR	704	96,8
	adenoma avançado	23	3,2
Topografia	cólon proximal	122	16,7
	cólon distal	265	36,5
	reto	334	45,9
	extra-cólon	81	11,1
	ignorado	6	0,8
Estadiamento clínico inicial	0	17	2,4
	I	114	16,2
	II	168	23,9
	III	209	29,7
	IV	157	22,3
	Ignorado	40	5,7
T	n.a.	23	
	is	16	2,3
	1	45	6,4
	2	99	14,1
	3	397	56,4
	4	66	9,4
	X	81	11,5
N	n.a.	23	
	0	296	42,0
	1	178	25,3
	2	95	13,5
	X	135	19,2
M	n.a.	23	
	0	515	73,2
	1	157	22,3
	X	32	4,5
Grau de diferenciação	n.a.	23	
	1	94	13,4
	2	515	73,2
	3	24	3,4
	ignorado	71	10,1
	n.a.	23	

Cont/Tabela 2

Características	Categoria	N	%
Tipo histológico	adenocarcinoma tubular/túbulo-viloso	233	32,0
	adenocarcinoma mucinoso/dif. mucinosa	104	14,3
	adenocarcinoma sem especificação	361	49,7
	adenoma avançado	23	3,2
	outros	5	0,7
Status atual da doença*	Vivo sem doença	416	57,2
	Vivo com doença	85	11,7
	Óbito por câncer	60	8,3
	Óbito outras causas	4	0,6
	Perda de seguimento	79	10,9
	Ignorado	83	11,4
História familiar	sim	222	30,5
	não	375	51,6
	Ignorado	130	17,8

n.a. não se aplica *informação colhida em agosto de 2012.

4.2 GENOTIPAGEM DE SNPs E TESTES DE ASSOCIAÇÃO

O número de indivíduos genotipados para cada *SNP* e as respectivas frequências genótípicas estão na Tabela 3:

Tabela 3 - Frequências genotípicas em casos e controles (Hospital A.C.Camargo – 2008-2012)

SNP	rs4939827 (SNP06)		rs10411210 (SNP09)		rs4444235 (SNP16)		rs9929218 (SNP21)		rs4779584 (SNP26)		rs961253 (SNP45)		rs10795668 (SNP59)		rs6983267 (SNP71)		rs3802842 (SNP82)		rs16892766 (SNP83)	
	casos	controles	casos	controles	casos	controles	casos	controles	casos	controles	casos	controles	casos	controles	casos	controles	casos	controles	casos	controles
Genótipos																				
Teste genotípico	p = 0,00013		p = 0,034		p = 0,020		p = 0,912		p = 0,193		p = 0,885		p = 0,842		p = 0,219		p = 0,024		p = 0,033	
Homozigoto comum	161 (22,1%)	250 (33,8%)	523 (71,9%)	492 (66,5%)	207 (28,5%)	212 (28,6%)	363 (49,9%)	382 (51,6%)	389 (53,5%)	378 (51,1%)	318 (43,7%)	317 (42,8%)	390 (53,6%)	395 (53,4%)	235 (32,3%)	279 (37,7%)	359 (49,4%)	427 (57,7%)	575 (79,1%)	643 (86,9%)
Homozigoto raro	187 (25,7%)	147 (19,9%)	31 (4,3%)	27 (3,6%)	174 (23,9%)	161 (21,8%)	61 (8,4%)	57 (7,7%)	52 (7,2%)	68 (9,2%)	85 (11,7%)	97 (13,1%)	62 (8,5%)	63 (8,5%)	123 (16,9%)	106 (14,3%)	66 (9,1%)	40 (5,4%)	10 (1,4%)	11 (1,5%)
Heterozigoto	336 (46,2%)	338 (45,7%)	167 (23,0%)	221 (29,9%)	341 (46,9%)	367 (49,6%)	295 (40,6%)	300 (40,5%)	252 (34,7%)	294 (39,7%)	310 (42,6%)	326 (44,1%)	261 (35,9%)	282 (38,1%)	334 (45,9%)	355 (48,0%)	266 (36,6%)	273 (36,9%)	115 (15,8%)	83 (11,2%)
Indeterminado	43 (5,9%)	5 (0,7%)	6 (0,8%)	0	5 (0,7%)	0	8 (1,1%)	1 (0,1%)	34 (4,7%)	0	14 (1,9%)	0	14 (1,9%)	0	35 (4,8%)	0	36 (4,9%)	0	27 (3,7%)	3 (0,4%)
Total	727	740	727	740	727	740	727	740	727	740	727	740	727	740	727	740	727	740	727	740

As frequências alélicas, o número total de alelos e suas frequências, a taxa de genotipagem e o teste de equilíbrio de Hardy-Weinberg (EHW) estão representados na tabela 4:

Tabela 4 - Frequências alélicas nos casos e controles (Hospital A.C.Camargo – 2008-2012).

	TESTE DE EHW		FREQUÊNCIA ALÉLICA				NÚMERO DE ALELOS				FREQUÊNCIA		TAXA DE GENOTIPAGEM		
	casos	controles	casos		controles		casos		controles		alelo major	alelo minor	casos	controles	Global
SNP	p	p	alelo major	alelo minor	alelo major	alelo minor	alelo major	alelo minor	alelo major	alelo minor	alelo major	alelo minor	%	%	%
rs4939827 (SNP06)	0,7018	0,0978	0,48	0,52	0,57	0,43	658	710	838	632	0,53	0,47	94,1%	99,3%	96,7%
rs10411210 (SNP09)	0,0007	0,7158	0,84	0,16	0,81	0,19	1213	229	1205	275	0,83	0,17	99,2%	100,0%	99,6%
rs4444235 (SNP16)	0,1565	0,9411	0,48	0,52	0,53	0,47	689	755	791	689	0,51	0,49	99,3%	100,0%	99,7%
rs9929218 (SNP21)	0,9279	0,9273	0,71	0,29	0,72	0,28	1021	417	1064	414	0,72	0,28	98,9%	99,9%	99,4%
rs4779584 (SNP26)	0,2319	0,3270	0,74	0,26	0,71	0,29	1030	356	1050	430	0,73	0,27	95,3%	100,0%	97,7%
rs961253 (SNP45)	0,5022	0,3750	0,66	0,34	0,65	0,35	946	480	960	520	0,66	0,34	98,1%	100,0%	99,0%
rs10795668 (SNP59)	0,0576	0,2311	0,73	0,27	0,72	0,28	1041	385	1072	408	0,73	0,27	98,1%	100,0%	99,0%
rs6983267 (SNP71)	0,8151	0,7557	0,58	0,42	0,62	0,38	804	580	913	567	0,60	0,40	95,2%	100,0%	97,6%
rs3802842 (SNP82)	0,1146	0,7613	0,71	0,29	0,76	0,24	984	398	1127	353	0,74	0,26	95,0%	100,0%	97,5%
rs16892766 (SNP83)	0,1287	0,0005	0,9	0,10	0,93	0,07	1265	135	1369	105	0,92	0,08	96,3%	99,6%	98,0%

Os testes de associação genética com os respectivos modelos estão representados na tabela 5:

Tabela 5 - Testes de associação genética (Hospital A.C.Camargo – 2008-2011).

	Modelo Alélico OR [95%I.C.]		P	Modelo Dominante OR [95%I.C.]		P	Modelo Recessivo OR [95%I.C.]		P	Modelo Aditivo OR [95%I.C.]		P
	alelo <i>major</i>	alelo <i>minor</i>		[HMZR + HTZ] x HMZC	HMZR x [HTZ + HMZC]		[HTZ] x [HMZC]			[HMZR] x [HTZ]		
SNP06	0.70 [0.60- 0.81]	1.43 [1.23- 1.66]	2.12 x 10 ⁻⁵	1.67 [1.33- 2.11]	0.0001	1.51 [1.18- 1.93]	0.011	1.54 [1.20- 1.98]	1.28 [0.98 – 1.67]	3.49 x 10 ⁻⁵		
SNP09	1.21 [1.00 – 1.47]	0.83 [0.68 – 1.00]	0.080	0.75 [0.60 – 0.94]	0.040	1.19 [0.70 – 2.01]	0.749	0.71 [0.56 – 0.90]	1.52 [0.87 – 2.64]	0.083		
SNP16	0.79 [0.69- 0.92]	1.26 [1.09- 1.45]	0.011	1.26 [1.00 – 1.60]	0.097	1.45 [1.14 – 1.83]	0.012	1.13 [0.88 – 1.45]	1.38 [1.07 – 1.78]	0.012		
SNP21	0.95 [0.81 – 1.12]	1.05 [0.89 – 1.23]	0.629	1.05 [0.85 – 1.29]	0.645	1.11 [0.76 – 1.62]	0.737	1.03 [0.83 – 1.28]	1.09 [0.73 – 1.62]	0.628		
SNP26	1.18 [1.00 – 1.40]	0.84 [0.72 – 1.00]	0.089	0.82 [0.66 – 1.00]	0.092	0.80 [0.55 – 1.17]	0.498	0.83 [0.67 – 1.04]	0.89 [0.60 – 1.33]	0.097		
SNP45	1.07 [0.92 – 1.24]	0.94 [0.80 – 1.09]	0.515	0.93 [0.76 – 1.15]	0.623	0.90 [0.66 – 1.22]	0.824	0.95 [0.76 – 1.18]	0.92 [0.66 – 1.28]	0.524		
SNP59	1.03 [0.87 – 1.21]	0.97 [0.83 – 1.14]	0.739	0.95 [0.77 – 1.17]	0.682	1.02 [0.71 – 1.48]	1	0.94 [0.75 – 1.17]	1.06 [0.72 – 1.57]	0.739		
SNP71	0.86 [0.74 – 1.00]	1.16 [1.00 – 1.35]	0.086	1.18 [0.95 – 1.46]	0.2	1.29 [0.97 – 1.72]	0.188	1.12 [0.89 – 1.40]	1.23 [0.91 – 1.66]	0.082		
SNP82	0.77 [0.66 – 0.92]	1.29 [1.09 – 1.53]	0.010	1.26 [1.02 – 1.56]	0.072	1.85 [1.23 – 2.78]	0.009	1.16 [0.93 – 1.44]	1.69 [1.10 – 2.60]	0.011		
SNP83	0.72 [0.55 – 0.94]	1.39 [1.07 – 1.82]	0.038	1.49 [1.11 – 1.99]	0.035	0.96 [0.40 – 2.27]	0.920	1.55 [1.14 – 2.10]	0.66 [0.27 – 1.62]	0.049		

HMZR = homozigotos raros; HMZC = homozigotos comuns; HTZ = heterozigotos

Dos dez *SNPs*, cinco (06, 09, 16, 82, 83) foram associados significativamente ($p \leq 0,05$) com o risco de CCR e dois (26 e 71) tenderam à associação ($p < 0,1$). O *SNP* 06 mostrou a associação mais significativa entre todos os *SNPs* em todos os modelos genéticos. O *SNP* 09 foi o único preditor de menor risco, principalmente no modelo dominante (diminuição de

25%), enquanto que os *SNP* 16 e 82 foram associados com maior risco no modelo recessivo (aumento de 45% e 85%, respectivamente). O *SNP* 83 conferiu maior risco principalmente no modelo dominante (aumento de quase 50%). Os *SNPs* 26 e 71, por sua vez, obtiveram uma associação marginalmente significativa somente nos modelos alélico e aditivo, com tendência ao aumento do risco com o *SNP* 71 e diminuição do risco com o *SNP* 26. Em suma, dos cinco *SNPs* associados com o risco de CCR, dois (*SNP* 16 e 82) conferiram maior risco entre os homozigotos raros do que entre os heterozigotos e homozigotos comuns juntos, evidenciado pelo modelo recessivo e três (06, 09 e 83) o efeito no risco foi maior entre os heterozigotos e homozigotos raros juntos do que entre os homozigotos comuns, evidenciado pelo modelo dominante.

Na tabela 6 estão representados os cinco *SNPs* associados ao CCR com os respectivos alelos ancestral e variante, as frequências do alelo de risco em comparação com europeus, a magnitude do efeito do alelo variante sobre o risco e o risco atribuível populacional correspondente ao alelo variante, que é definido como a redução na incidência da doença se a população não fosse exposta ao alelo de risco.

Tabela 6 - Os cinco *SNPs* associados com o risco de CCR (Hospital A.C.Camargo – 2008-2011).

<i>SNP</i>	alelo ancestral*	alelo variante (risco)	frequência do alelo de risco (FAR)	frequência do alelo <i>minor</i> (europeus)*	magnitude do efeito	risco atribuível populacional** (%)
rs4939827 (SNP 06)	<i>T (major)</i>	<i>C (minor)</i>	0,47	0,47	umenta 43%	16,8
rs10411210 (SNP 09)	<i>C (major)</i>	<i>T (minor)</i>	0,17	0,10	diminui 25%	---
rs4444235 (SNP 16)	<i>T (major)</i>	<i>C (minor)</i>	0,49	0,48	umenta 26%	11,3
rs3802842 (SNP 82)	<i>C (minor)</i>	<i>C (minor)</i>	0,26	0,27	umenta 29%	7,0
rs16892766 (SNP 83)	<i>A (major)</i>	<i>C (minor)</i>	0,08	0,09	umenta 39%	3,0

*segundo dados do 1000 Genomes (www.1000genomes.org) **calculado pela fórmula: $100(\text{FAR}(\text{OR} - 1))/(1 + \text{FAR}(\text{OR} - 1))$

Fonte: TENESA e DUNLOP (2009)

4.3 CORRELAÇÕES GENÓTIPO-FENÓTIPO

Na Tabela 7 são apresentadas as correlações estatisticamente significativas das variáveis clínico-patológica-familiais com os genótipos:

Tabela 7 - Correlações genótipo-fenótipo (Hospital A.C.Camargo – 2008-2011).

Variável	Categoria	SNP	Genótipo	OR [I.C. 95%]		p
				Bethesda	Esporádico	
Classificação CCR	Bethesda	82	AC	1.60 [1.14 – 2.24]	0.62 [0.45 – 0.87]	0.006
	esporádico		CC	0.62 [0.44 – 0.86]	1.62 [1.16 – 2.26]	0.004
			TT	0.43 [0.20 – 0.95]	2.30 [1.06 – 5.01]	0.023
			TT	0.65 [0.45 – 0.95]	1.54 [1.05 – 2.24]	0.023
Idade	< 50 anos	26	TT	0.43 [0.17 – 1.11]	2.95 [1.15 – 7.56]	0.010
	≥ 50 anos	82	AC	1.51 [1.05 – 2.17]	0.67 [0.47 – 0.96]	0.030
Sexo	masculino	45	AA	1.76 [1.10 – 2.82]	0.57 [0.35 – 0.91]	0.016
	feminino					
Tipo de tumor	adenoma	26	TT	3.68 [1.31 – 10.36]		0.028
	cólon proximal	45	CC	0.60 [0.40 – 0.91]		0.014
Local do tumor	cólon distal	83	CC	7.02 [1.48 – 33.32]		0.005
	reto		AC	1.41 [1.02 – 1.93]		0.035
			CC	1.37 [1.01 – 1.85]		0.040
			AC	0.70 [0.51 – 0.95]		0.022
		45	AC	0.71 [0.52 – 0.95]		0.023
		CC	1.34 [1.00 – 1.81]		0.053	

Cont/ Tabela 7

	0	16	CC	2.83 [1.07 – 7.46]	0.025
			CT	0.34 [0.11 – 1.04]	0.041
		26	CT	0.24 [0.05 – 1.06]	0.026
			CC	3.41 [0.96 – 12.07]	0.035
Estadiamento	I	59	GG	2.48 [1.34 – 4.60]	0.006
clínico		83	CC	7.46 [2.07 – 26.88]	0.002
	II	71	GT	1.56 [1.08 – 2.24]	0.017
	III	--	--	--	n.s.
	IV	06	CC	0.63 [0.41 – 0.98]	0.035
	bem diferenciado	--	--	--	n.s.
Grau de	mod.diferenciado	83	CC	3.71 [0.97 – 5.92]	0.043
diferenciação	pouco	06	CT	2.40 [0.97 – 5.92]	0.048
	diferenciado				
Tipo	tubular/túbulo-	09	CC	0.60 [0.36 – 1.00]	0.045
histológico	viloso				
	mucinoso/dif.	--	--	--	n.s.
	mucinosa				
História	sim	06	TT	0.58 [0.38 – 0.91]	0.014
familiar	não				
	vivo com doença	--	--	--	n.s.
	vivo sem doença	06	CC	1.49 [1.02 – 2.17]	0.035
Status da		71	GG	0.70 [0.50 – 0.99]	0.042
doença			GT	1.39 [1.01 – 1.93]	0.045
	óbito por câncer	09	TT	2.98 [1.15 – 7.70]	0.040
		45	CC	2.11 [1.23 – 3.62]	0.006
			AC	0.57 [0.33 – 1.00]	0.045

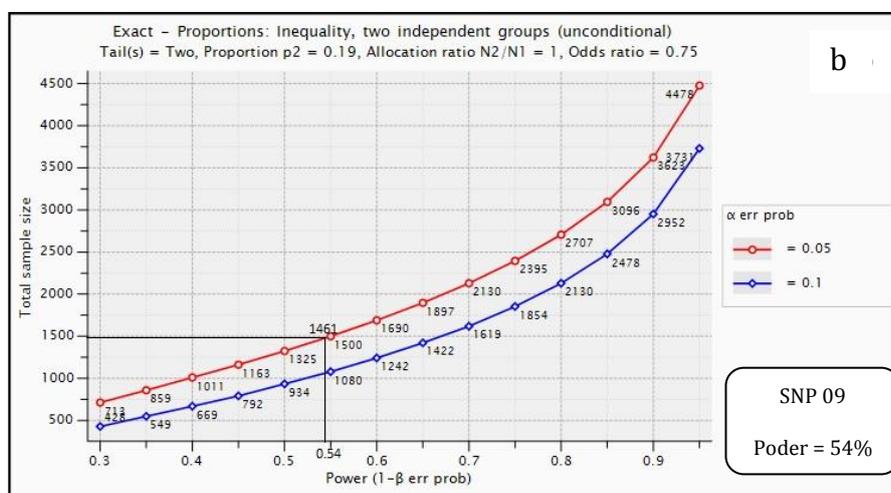
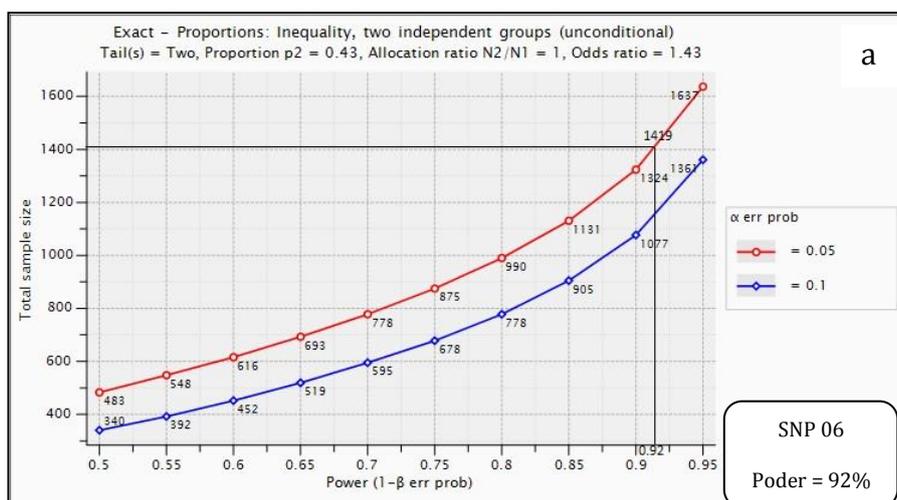
Quanto à classificação do CCR, os heterozigotos para o *SNP* 82 predisseram maior risco para o tipo Bethesda, enquanto que os homozigotos comuns para o *SNP* 82, os homozigotos raros para o *SNP* 26 e 16 tiveram maior risco para o tipo esporádico; em relação à idade ao diagnóstico, os heterozigotos para o *SNP* 82 predisseram maior risco para os menores de 50 anos, enquanto que os homozigotos raros para o *SNP* 26 predisseram maior risco ao grupo mais velho; quanto ao sexo, os homozigotos comuns para o *SNP* 45 predisseram maior risco entre os homens e menor entre as mulheres; já em relação ao tipo de tumor, os homozigotos raros para o *SNP* 26 predisseram maior risco para o desenvolvimento de adenomas; considerando a localização do tumor, os homozigotos raros para o *SNP* 83 e os heterozigotos para o *SNP* 82 predisseram maior risco para os tumores do cólon distal, enquanto que para os do cólon proximal, os homozigotos raros para o *SNP* 45 tiveram menor risco. Já para os para tumores do reto, os homozigotos raros para o *SNP* 45 e os homozigotos comuns para o *SNP* 82 tiveram maior risco, enquanto que os heterozigotos para esses dois *SNPs* tiveram menor risco. Em relação ao estadiamento clínico inicial, o estágio 0 foi correlacionado com os homozigotos raros para o *SNP* 16 e homozigotos comuns para o *SNP* 26; o estágio I, com os homozigotos comuns para o *SNP* 59 e raros para o *SNP* 83; o estágio II, como os heterozigotos para o *SNP* 71 e o estágio IV com menor risco entre os homozigotos raros para o *SNP* 06. Não houve correlação estatisticamente significativa entre nenhum dos genótipos e o estágio III. Em relação ao grau de diferenciação, os heterozigotos para os *SNP* 06 correlacionaram-se com os tumores pouco

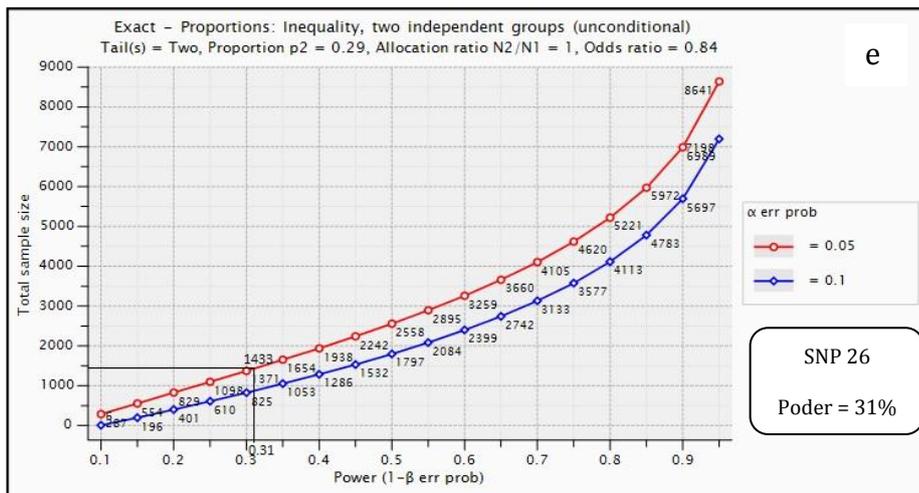
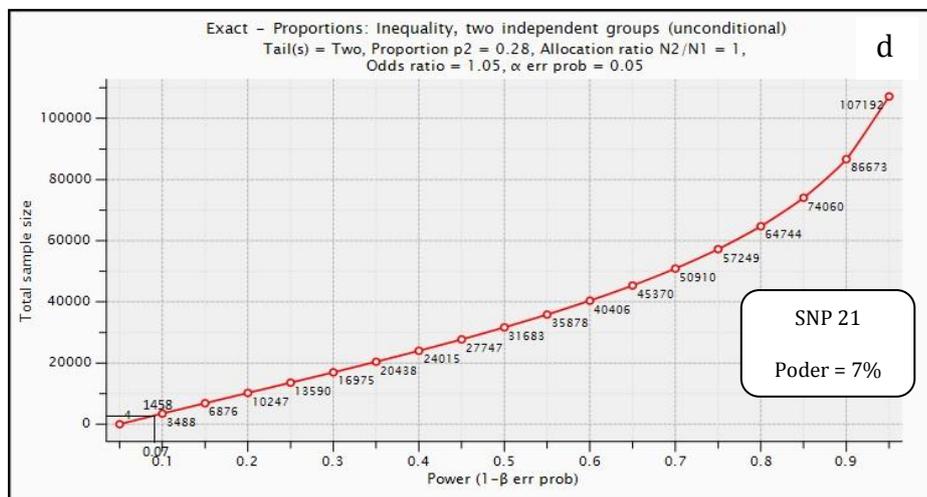
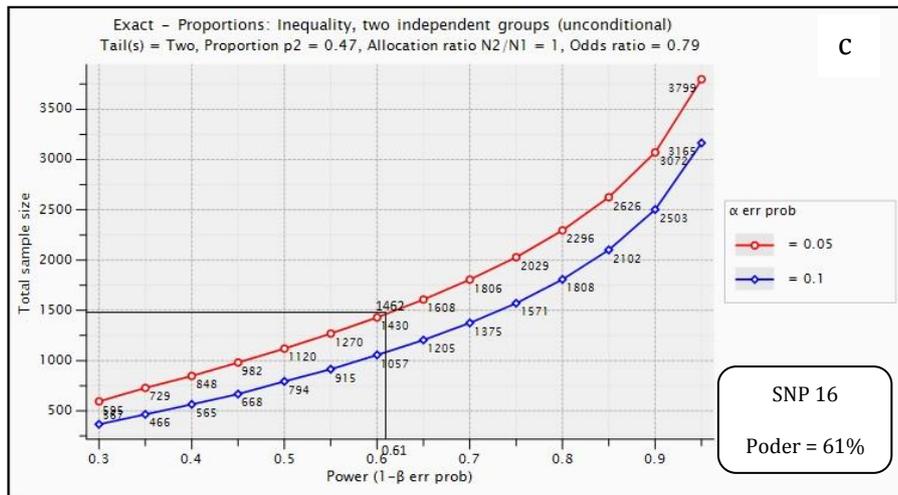
diferenciados, enquanto que os moderadamente diferenciados com os homozigotos raros para o *SNP* 83. Os bem diferenciados não obtiveram significância estatística. Quanto ao tipo histológico, os homozigotos raros para o *SNP* 09 tenderam a um menor risco para o tipo tubular. A presença de um parente até 3º grau com CCR não foi correlacionada com nenhum dos genótipos, sendo que os homozigotos comuns para o *SNP* 06 apontaram para uma menor chance de história familiar. Em relação ao status da doença, os homozigotos raros para os *SNPs* 09 e 45 predisseram maior risco para óbito devido ao CCR, enquanto que os homozigotos raros para o *SNP* 06 e os heterozigotos para o *SNP* 71 predisseram maior chance de permanecerem vivos sem doença. Não houve correlação estatisticamente significativa entre os genótipos e a chance de permanecer vivo com doença.

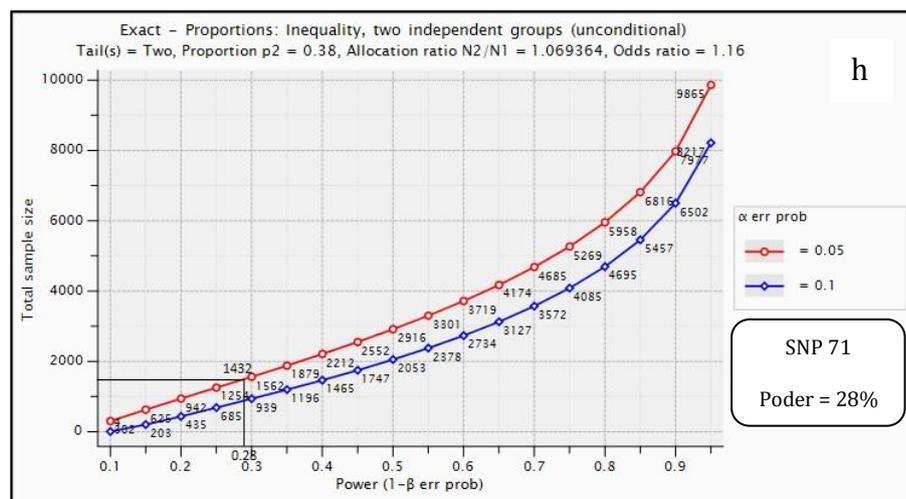
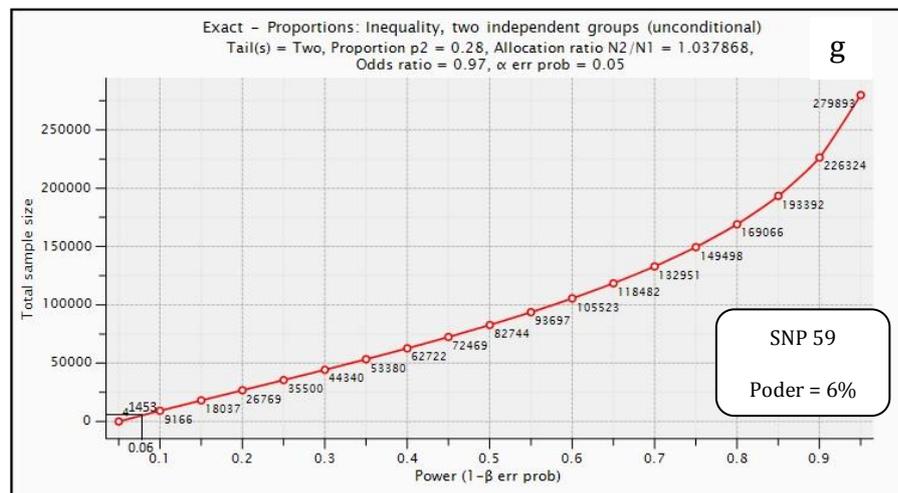
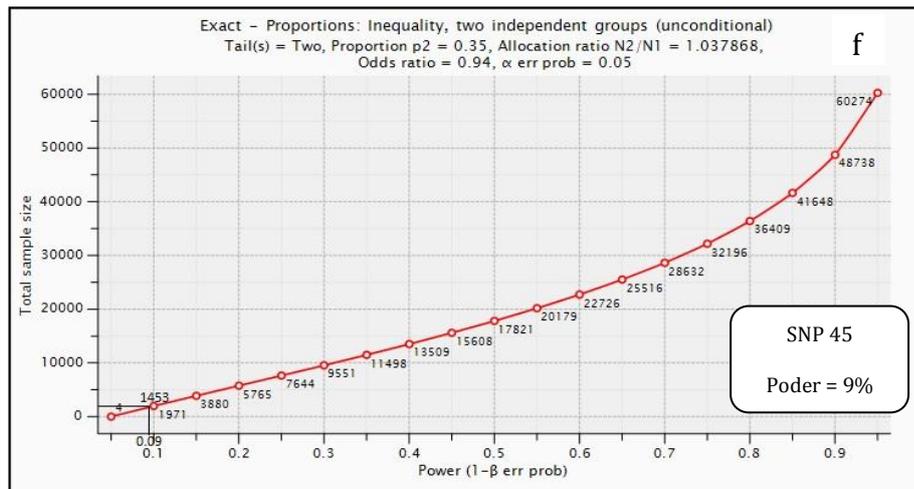
4.4 TAMANHO DA AMOSTRA E PODER ESTATÍSTICO

Inicialmente, o presente estudo havia sido desenhado com um tamanho amostral de 700 indivíduos (350 casos e 350 controles). Contudo, a *posteriori*, utilizando um programa de cálculo de tamanho de amostra (G*POWER 3.1, FAUL et al. 2009), estimou-se um total de 1.500 indivíduos (750 casos e 750 controles) para identificar uma associação entre os dez *SNPs*, com um *odds ratio* estimado de 1.5, poder estatístico de 80% e α bicaudado de 0,005 (corrigido pelo método de Bonferroni). Ao calcular o poder estatístico para cada *SNP*, com suas respectivas frequências alélicas (alelo *minor*) entre os casos e controles estudados, utilizando o mesmo

programa, o único *SNP* que atingiu poder superior a 80% foi o *SNP* 06 (α bicaudado de 0,05). Abaixo são apresentados os gráficos dos cálculos do poder estatístico para cada um dos *SNPs* (Figura 8 a-j).







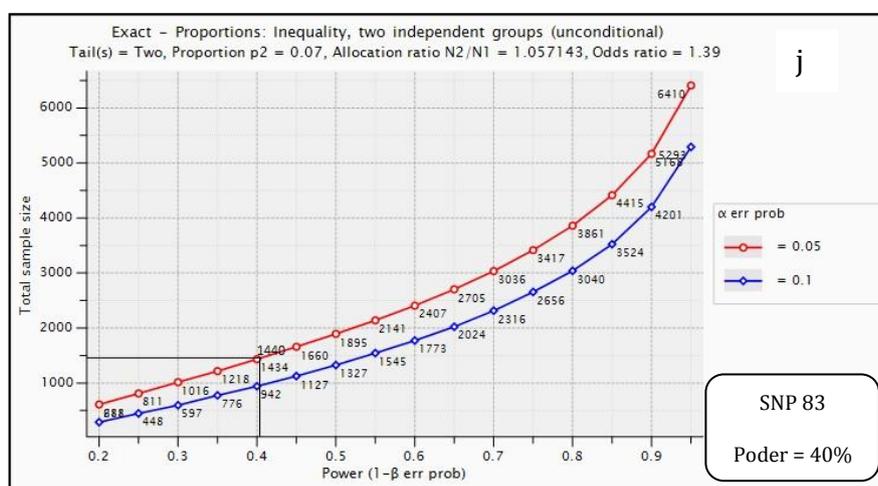
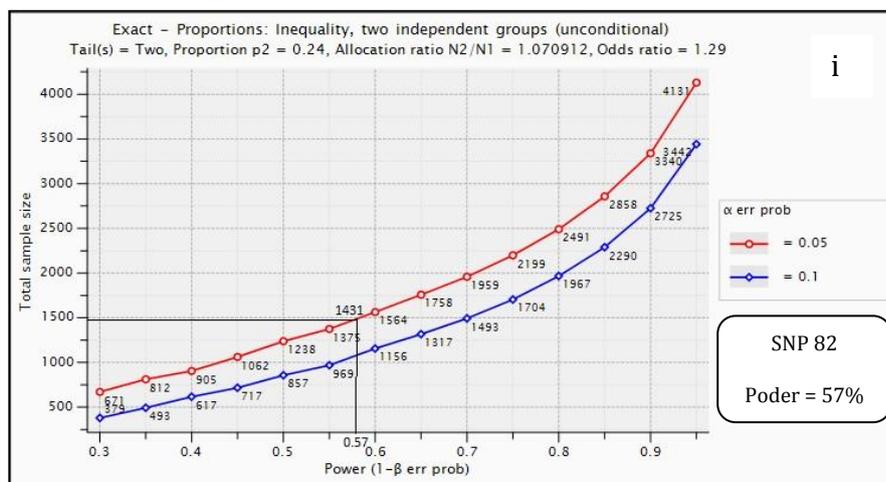


Figura 8 - A – J - Cálculos do poder estatístico (linha vermelha: $\alpha = 0,05$; linha azul: $\alpha = 0,1$)

5 DISCUSSÃO

5.1 PODER ESTATÍSTICO

O poder de um teste estatístico é a probabilidade do teste resultar em um resultado estatisticamente significativo dado que a hipótese nula seja falsa, ou seja, a chance de que um estudo detecte um efeito verdadeiro. Vários fatores podem afetar o poder para detectar uma associação entre um fator de risco e uma doença, principalmente: a magnitude do efeito do fator de risco, o tamanho da amostra, a prevalência da doença, o desenho do estudo e a taxa específica de falsos positivos (EVANS e PURCELL 2007). Na ocasião de um estudo genético caso-controle para testar a associação entre um locus/loci ou *SNP*(s) e uma doença, determinados fatores podem influenciar o poder: o efeito sobre o risco, o modo de herança, a razão entre casos e controles, as frequências alélicas e a extensão do desequilíbrio de ligação entre o marcador e o *SNP* funcional. O pesquisador deve decidir antecipadamente quais parâmetros levará em conta, uns sob seu controle (tamanho da amostra, taxa de erro tipo I ou α , prevalência da doença) e outros não, como as frequências alélicas, o grau de desequilíbrio de ligação, o modo de herança e o efeito genético. Em geral, o poder para detectar uma associação é diretamente proporcional à frequência do alelo da doença (*minor*) e à magnitude do efeito, embora dependa, em parte, do modo de herança. Por exemplo, na herança dominante, mais indivíduos serão

necessários se a frequência do alelo da doença for alta. Basicamente, há duas maneiras de aumentar o poder de um teste estatístico: aumentando o tamanho da amostra e a magnitude do efeito, neste último, ao selecionar acuradamente os fenótipos ou genotipar mais densamente uma região de interesse. Por exemplo, selecionar pacientes com história familiar de câncer em parentes de 1º grau enriqueceria a contribuição genética e poderia reduzir o tamanho necessário da amostra em até quatro vezes (ANTONIOU e EASTON 2003). Similarmente, o poder estatístico pode ser aumentado com a inclusão de “super-controles”, como indivíduos idosos saudáveis rastreados com colonoscopia, ou aqueles sem história familiar de CCR (TENESA e DUNLOP 2009). Da mesma forma, realizar uma metanálise, onde se aumenta o número de casos e controles por estudos replicativos independentes em múltiplas fases em populações semelhantes eleva a significância estatística (HOULSTON et al. 2008).

Haveria pelo menos três explicações plausíveis interdependentes para a falta de associação no presente estudo:

1. poder limitado pela amostra insuficiente;
2. diferenças entre os padrões de desequilíbrio de ligação (DL) das variantes pesquisadas com as variantes causais entre diferentes grupos étnicos da população brasileira;
3. diferenças de penetrância das variantes causais e, conseqüentemente, dos efeitos dos alelos de risco.

A falta de associação estatisticamente significativa entre os *SNPs* 21, 26, 45, 59 e 71, embora houvesse uma tendência à associação com os

SNPs 26 e 71, muito provavelmente deveu-se à falta de poder pelo tamanho insuficiente da amostra. Assim, o número amostral deveria ser aumentado, levando em consideração tanto a frequência quanto a penetrância dos alelos de risco entre os casos e controles para evidenciar tal associação. Segundo o programa G*POWER 3.1, no caso do *SNP* 26, seria necessária uma amostra de cerca de 5.200 indivíduos para atingir um poder de 80% e α bicaudado de 5% (Figura 8e); o *SNP* 71, por sua vez, a genotipagem de cerca de 5.900 indivíduos atingiria tal poder (Figura 8h).

HE et al. (2011) falharam em replicar a associação em oito dos 11 *SNPs* pesquisados (dez neste estudo) em 2.705 indivíduos americanos-europeus (1.171 casos e 1.534 controles). Dos oito *SNPs* não associados no estudo norte-americano, quatro foram associados neste estudo (*SNPs* 06, 09 16 e 83) e três não foram associados com o risco (*SNPs* 21, 45 e 59). O poder do estudo para detectar os efeitos desses oito *SNPs* relatados nos GWAS foi 84% para o *SNP* 06, o único acima de 80% (assim como neste estudo) e entre 34% e 62% para os outros sete *SNPs*, denotando, assim, a falta de poder estatístico.

VON HOLST et al. (2010) também não conseguiram replicar a associação em seis dos 11 *SNPs* pesquisados (dez neste estudo) em 3.535 suecos (1.786 casos e 1.749 controles). Dos seis *SNPs* não associados, três foram associados neste estudo (*SNPs* 09 16, 82) e dois não foram associados (*SNPs* 21 e 45). Os autores argumentam que tal falha poderia ser explicada por populações diferentes, tamanho da amostra e métodos de genotipagem, porém as frequências genotípicas foram similares às dos

GWAS e o método empregado (TaqMan) foi validado. Outros estudos replicativos europeus verificaram associação com os *SNPs* 06, 59 e 71 (POYNTER et al. 2007; THOMPSON et al. 2009; MIDDELDORP et al. 2009).

Uma metanálise recente utilizou os dados de GWAS com 6.322 indivíduos americanos (2.906 casos e 3.416 controles) e replicou em mais 17.262 indivíduos (8.161 casos e 9.101 controles) de dez estudos independentes de vários países (Canadá, Alemanha, França, Israel e Estados Unidos), totalizando uma amostra de 23.584 indivíduos (11.067 casos e 12.517 controles). O estudo encontrou associação em oito dos dez *SNPs* pesquisados (os mesmos deste estudo) com um nível de significância de 0,05; porém, nenhum dos 10 atingiu significância estatística após ajustar para múltiplas análises. Os autores argumentaram que a falha em replicar os GWAS possa ter devido à metodologia aplicada, já que o tamanho da amostra foi suficientemente grande e a população selecionada foi de indivíduos “brancos hispânicos”, o que diminuiu a chance de diferentes DL (PETERS et al. 2012).

Portanto, a falta de replicação das associações encontradas nos GWAS pode dever-se a três razões principais (STRACHAN e READ 2011):

1. pareamento inadequado entre casos e controles: de fato, esta é uma preocupação importante, até mesmo nos estudos mais cuidadosamente desenhados;
2. correção insuficiente por múltiplos testes: no intuito de evitar o erro tipo I;

3. poder insuficiente: um estudo com baixo poder estatístico pode ocasionalmente ter a sorte de encontrar uma associação verdadeira, porém, tal sorte é improvável num estudo semelhante, já que os estudos replicativos precisam de muito mais poder do que o estudo inicial.

5.2 FREQUENCIAS E EFEITOS DOS SNPs

Nas doenças comuns, como no CCR, estima-se que grande parte do risco genético deva-se à herança de múltiplos *loci* seguindo um modelo poligênico, cada um com frequências alélicas comuns (FAM maior do que 5%), e efeitos de baixa magnitude, com riscos relativos entre 1,0 e 1,5 (TOMLINSON et al. 2010). Desse modo, para se detectar tais efeitos pequenos, é necessário que a amostra seja suficientemente grande. Esta estratégia foi claramente validada por metanálises dos dados dos GWAS europeus, onde dezenas de milhares de indivíduos foram genotipados por plataformas de grande cobertura do genoma, seguido de validação por múltiplas fases, com séries independentes de caso e controles. Mesmo assim, somente cerca de 20 SNPs comuns com efeitos modestos sobre o risco (menores do que 30%) foram detectados até o momento, cada um com um nível de significância abaixo de um limiar pré-definido controlado por múltiplos testes ($5,0 \times 10^{-8}$). Na tabela 8 estão representados os GWAS europeus com os dez SNPs de suscetibilidade ao CCR que foram investigados neste estudo:

Tabela 8 - Dados dos dez *SNPs* identificados pelos *GWAS* em populações europeias.

<i>SNP</i>	alelo de risco	frequência do alelo de risco	magnitude do efeito (OR alélico)	risco atribuível populacional (%)	tamanho total da amostra	referências
rs4939827 (SNP 06)	<i>T</i> (major)	0,52	aumenta 18%	8,6	15.362	<i>Broderick et al. 2007</i>
rs10411210 (SNP 09)	<i>T</i> (minor)	0,10	diminui 13%	--	41.259	<i>Houlston et al. 2008</i>
	<i>C</i> (major)	0,90	aumenta 15%	11,9		
rs4444235 (SNP 16)	<i>C</i> (minor)	0,46	aumenta 11%	4,8	41.259	<i>Houlston et al. 2008</i>
rs9929218 (SNP 21)	<i>G</i> (major)	0,71	aumenta 10%	6,6	41.259	<i>Houlston et al. 2008</i>
rs4779584 (SNP 26)	<i>T</i> (minor)	0,18	aumenta 26%	4,5	14.663	<i>Jaeger et al. 2008</i>
rs961253 (SNP 45)	<i>A</i> (minor)	0,36	aumenta 12%	4,0	41.259	<i>Houlston et al. 2008</i>
rs10795668 (SNP 59)	<i>G</i> (major)	0,67	aumenta 12%	7,4	37.371	<i>Tomlinson et al. 2008</i>
rs6983267 (SNP 71)	<i>G</i> (major)	0,51	aumenta 21%	9,7	14.470	<i>Tomlinson et al. 2007</i>
rs3802842 (SNP 82)	<i>C</i> (minor)	0,29	aumenta 12%	3,4	27.794	<i>Tenesa et al. 2008</i>
rs16892766 (SNP 83)	<i>C</i> (minor)	0,07	aumenta 25%	1,7	37.371	<i>Tomlinson et al. 2008</i>

No presente estudo houve associação com o risco de CCR em metade dos *SNPs* pesquisados (06, 09 16, 82 e 83), cujos alelos de risco mostraram frequências próximas às dos *GWAS* europeus (8% a 49% neste estudo; 7% a 67% nos *GWAS* europeus), com exceção do alelo de risco do *SNP* 06 (*C* minor=47% neste estudo e *T* major=52% nos *GWAS*). Os efeitos também foram modestos (21% a 43% neste estudo; 10 a 26% nos *GWAS*). O *SNP* 06 foi a variante que obteve o maior efeito sobre o risco (43%) e com a maior associação ($p_{\text{trend}} = 3,49 \times 10^{-5}$), conferindo o maior risco populacional (16,8%), enquanto que no *GWAS* europeu aumentou-se o risco

em até 23%, representando 8,6% do risco populacional (TENESA e DUNLOP 2009). No estudo original, também foi o *SNP* que mostrou maior associação ($p_{\text{trend}} = 1,0 \times 10^{-12}$) (BRODERICK et al. 2007). O *SNP* 09 (rs10411210) foi identificado em associação com um risco diminuído de maneira dose-dependente (em heterozigotos diminui até 17% e em homozigotos, até 41%) (HOULSTON et al. 2008). Neste estudo, porém, o efeito foi observado apenas no modelo dominante, onde os homozigotos raros e os heterozigotos possuem o mesmo risco. Vale ressaltar que nos estudos europeus o alelo *major(C)* confere aumento de 15% do risco, representado 12% do risco populacional (TENESA e DUNLOP 2009). Neste estudo houve uma tendência ao maior risco (1,0 a 1,47), porém não foi estatisticamente significativo ($p = 0,08$). O *SNP* 16, por sua vez, foi associado um risco maior neste estudo (26%) do que nos GWAS europeus (11%), assim como o risco populacional (11,3% contra 4,8%) (TENESA e DUNLOP 2009). Da mesma forma, os *SNP* 82 e 83 também aumentaram mais o risco neste estudo (29% e 39%, respectivamente) do que nos GWAS (12% e 25%), assim como o risco populacional (7% e 3%, respectivamente neste estudo contra 3,4% e 1,7%, nos GWAS) (TENESA e DUNLOP 2009).

5.3 ESTRATIFICAÇÃO POR ANCESTRALIDADE

Com a mesma relevância, outro fator que se deve levar em consideração é a estratificação populacional por ancestralidade como fator de confusão. As comparações entre grupos étnicos diferentes têm

demonstrado que as frequências alélicas podem variar amplamente, assim, alguns *SNPs* podem ser informativos em uma população e não em outras, com efeitos maiores ou menores, dependendo da ancestralidade. Esta é uma das possíveis explicações para algumas associações falso-positivas relatadas na literatura (TOMLINSON et al. 2010). O mapeamento de miscigenação (do inglês, *admixture mapping*) é desenhado para avaliar a variação da ancestralidade, utilizando painéis de marcadores informativos de ancestralidade (do inglês, *ancestry-informative markers*), que permitem comparar o DL da miscigenação com o DL padrão das populações parentais. A distribuição das frequências haplotípicas na população miscigenada pode diferir substancialmente das populações parentais. Em um dado *locus*, as frequências alélicas podem variar entre as populações parentais, assim, a frequência alélica esperada na população miscigenada é a combinação linear das frequências alélicas das populações parentais com a proporção de miscigenação da amostra. Isto, ultimamente, vai determinar quantos marcadores são necessários para o mapeamento das variantes funcionais, reduzindo as taxas de erros por falsos positivos ao eliminar o fator de confusão por ancestralidade (SHRINER et al. 2011).

Aproximadamente 90% dos *GWAS* foram conduzidos com a utilização de amostras de indivíduos de ancestralidade europeia, isto é, numa população geneticamente homogênea de ambientes relativamente homogêneos (NEED e GOLDSTEIN 2009).

Um exemplo de doença vastamente estudada na população europeia por mais de 20 *GWAS* é o diabetes tipo 2, os quais identificaram cerca de 50

SNPs significativos, e cuja metanálise replicou os resultados em mais de 94.000 amostras (VISSCHER et al. 2012). Mais recentemente, estudos utilizaram amostras de asiáticos orientais, sul-asiáticos e hispânicos e mostraram que, apesar das diferenças entre as frequências alélicas e os padrões de DL, a maioria dos *SNPs* encontrados em um grupo étnico mostrou evidência de associação em outro grupo. Isto indica que as variantes comuns detectadas pelos *GWAS* são, provavelmente, o resultado de alelos amplamente distribuídos que são altamente frequentes nas populações, refletindo que há um alto DL de uma variante causal com pequeno efeito (VISSCHER et al. 2012).

Da mesma forma, a vasta maioria dos *GWAS* em câncer de mama e próstata foi realizada em populações europeias. Recentemente, estudou-se a validação de *SNPs* de risco para cânceres de próstata e mama em duas populações, afro-americana e hispano-americana. Para avaliar a estratificação e ajustar pela ancestralidade, foram utilizados 200 marcadores informativos de ancestralidade, baseando-se nas frequências alélicas ancestrais das populações africana e europeia do projeto *HapMap*. O estudo verificou que alguns *SNPs* permaneceram significativamente associados ao risco enquanto que outros não, após o ajuste para estratificação e para o efeito da interação com o *SNP* de interesse. Os autores concluíram que a estratificação populacional pode causar inflação ou deflação da significância estatística dos *SNPs* identificados pelos *GWAS*, e que o não-pareamento perfeito entre casos e controles das frequências alélicas de cada grupo racial pode revelar associações falso-positivas. Por exemplo, de todos os *SNPs*

estudados para câncer de próstata, somente um (rs16901979) foi significativo nos afro-americanos após o ajuste (BARNHOLTZ-SLOAN et al. 2011).

Igualmente às outras doenças complexas e outros tumores, a maioria dos GWAS em CCR foi empregada nas populações europeias, com vários estudos replicativos nos próprios europeus, que encontraram associações estatisticamente significativas com os SNPs dos estudos originais (POYNTER et al. 2007; TUUPANEN et al. 2008; BERNDT et al. 2008; THOMPSON et al. 2009; MIDDELDORP et al. 2009; NIITTYMAKI et al. 2010).

Poucos estudos foram replicados em diferentes etnias, como a população japonesa, chinesa e norte-americana (CURTIN et al. 2009; MATSUO et al. 2009; KUPFER et al. 2009; XIONG et al. 2010; HE et al. 2011).

HE et al. (2011) estudaram, nos Estados Unidos 11 SNPs associados ao risco de CCR em 3.311 casos e 4.466 controles de cinco etnias auto-relatadas (americanos europeus, americanos africanos, havaianos nativos, americanos japoneses e latino-americanos). Para avaliar os potenciais efeitos de confusão da estratificação populacional, foi realizada uma análise de componente principal (do inglês, *principal component analysis*) da genotipagem de 1.389 SNPs entre casos e controles de uma corte multi-étnica independente. As frequências alélicas variaram substancialmente entre os grupos étnicos. Nos americanos europeus, somente três SNPs foram associados com significância estatística (rs4779584 ou SNP26,

rs6983267 ou *SNP71* e rs3802842 ou *SNP82*), enquanto que dois *SNPs* (rs961253 ou *SNP45* e rs6983267 ou *SNP71*) foram associados com maior risco em americanos africanos. Além disso, houve um efeito cumulativo estatisticamente significativo dos alelos de risco de CCR em todas as etnias estudadas, exceto nos americanos japoneses.

A variante rs6983267 (*SNP 71*) aumentou o risco em todas as populações estudadas: europeias, asiáticas e em cinco etnias norte-americanas, podendo ser caracterizado como um marcador pan-étnico de risco ao CCR (HE et al. 2011).

A identificação das variantes causais a partir de um *tagSNP* estatisticamente associado ao CCR é difícil, pois geralmente mais de um gene situa-se no bloco de DL indicado pela associação com o *tagSNP*. Por isso, estudar diferentes grupos étnicos com diferentes padrões de blocos de DL pode ser útil em refinar o locus de suscetibilidade antes do resequenciamento e análises funcionais. Um exemplo foi um estudo que identificou novas variantes próximas aos genes da via BMP: *GREM1*, *BMP4* e *BMP2* (TOMLINSON et al. 2011). A utilidade dessa abordagem depende do tamanho da amostra, da seleção dos fenótipos e das frequências dos alelos de risco em cada grupo étnico.

Assim, é mister expandir os *GWAS* para populações mais heterogêneas por três motivos principais:

- identificar novos *loci* que não foram identificados em populações europeias devido à uma baixa frequência alélica, resultando em um baixo poder estatístico;

- estabelecer a extrapolação ou replicação dos achados em populações europeias em populações não-europeias; e
- estudar doenças pouco prevalentes na Europa.

Neste estudo não foi investigada a estratificação populacional por ancestralidade, uma vez que um dos objetivos do estudo multicêntrico no qual este se insere, o Estudo Genético de Câncer Colorretal na Hispânia e nas Américas, é a busca de novos *SNPs* através da combinação do *GWAS* com o mapeamento de miscigenação entre as populações europeia e latino-americana, incluindo a brasileira.

Sabe-se, no entanto, que a população brasileira, embora ricamente miscigenada, tem uma maior prevalência de indivíduos de ascendência europeia, majoritariamente nas regiões Sul (79,5%) e Sudeste (74,2%) do país. PENA et al. (2011) utilizaram um painel com 40 marcadores informativos de ancestralidade (MIA) europeia, africana e ameríndia para estudar a heterogeneidade étnica no Brasil em quase mil indivíduos. Verificou-se que há discrepâncias entre o critério de cor auto-referida (branco, pardo e negro) e a ancestralidade. Por isso, preferiu-se não utilizar cor ou raça como variáveis para inferir ancestralidade. GIOLO et al. (2012) estudaram a ancestralidade da população do Sudeste brasileiro ao comparar 138 indivíduos brasileiros com os dados de 991 indivíduos de 11 populações do projeto *HapMap*, que incluem asiáticos, europeus, africanos e mexicanos. Foram utilizados três painéis de MIA e, após as análises de componente principal, concluiu-se que a população desta região do país é derivada da

miscigenação entre europeus, africanos e ameríndios, principalmente entre os dois primeiros.

A amostra deste estudo foi selecionada de um único hospital, majoritariamente de cunho privado, situado em um bairro central de São Paulo, cuja maioria dos pacientes referiu ser descendente de europeus. Como as frequências alélicas dos *SNPs* assemelharam-se às da população europeia (segundo o estudo *1.000 Genomes*), inferir-se-ia que as frequências haplotípicas (entre o *tagSNP* e a variante causal) também o são, devido a um DL menos heterogêneo, como o da população europeia. Entretanto, com o intuito de replicar integralmente os dados dos *GWAS* europeus na população brasileira, assim como em outras populações recentemente miscigenadas, dever-se-ia controlar os fatores passíveis de causarem associações espúrias, como a ancestralidade, assim sendo, portanto, imprescindível analisar tal estratificação populacional.

5.4 CORRELAÇÕES GENÓTIPO-FENÓTIPO

Com o intuito de averiguar se algum dos 30 genótipos dos *SNPs* estudados seria capaz de predizer algum fenótipo, sejam características do tumor ou características dos pacientes, assim como a história familiar, confrontamos cada uma das variáveis dependentes, separadamente, com os 30 genótipos. O *SNP* 06 foi preditor de estar vivo sem doença (entre os homozigotos raros), sem história familiar (entre os homozigotos comuns), de tumores pouco diferenciados, e sem metástases à distância (estádio IV).

Um estudo recente analisou a significância prognóstica de 11 *SNPs* (cinco do presente estudo) associados ao CCR em 583 pacientes europeus-americanos e em africanos-americanos com diferentes estadiamentos, submetidos ou não à quimioterapia. Após as correções estatísticas e análises multivariadas (ajustada por sexo, estadio, etnia e tratamento), nenhuma das variantes mostrou alterar a sobrevida global. Entretanto, a variante rs4939827 (*SNP* 06) em homozigose mostrou correlação com estádios iniciais precoces ao diagnóstico (HOSKINS et al. 2012). No presente estudo, observou-se um menor risco ao estadio IV e maior chance de sobreviver sem doença entre os homozigotos raros do *SNP* 06 (genótipo CC), muito provavelmente por serem tratados com intenção curativa. Este *SNP* situa-se no íntron 3 do gene *SMAD7*, que é um antagonista da via da superfamília do TGF- β , bloqueando a ativação de sinais *downstream*. Distúrbios na expressão do *SMAD7* têm sido implicados na progressão do CCR (BRODERICK et al. 2007). A variabilidade interindividual da expressão de genes implicados em fatores de crescimento, apoptose e imunológicos pode influenciar mais os estágios tardios do que os eventos iniciais do CCR (TOMLINSON et al. 2010).

Os homozigotos raros do *SNP* 16 (genótipo CC) e do *SNP* 83 (genótipo CC) também se correlacionaram com um estadiamento mais precoce (0 e I, respectivamente). Interessantemente, parece que tais alelos parecem ter um efeito protetor quanto à progressão do CCR, apesar de aumentarem a suscetibilidade. A variante rs4444235 (*SNP* 16) situa-se próximo ao gene *BMP4*, outro membro da família TGF- β , que tem papel

importante na tumorigênese do CCR (WHITELAW et al. 1997). Um estudo espanhol, no entanto, encontrou associação do SNPs 83 e 71 com tumores em estágios mais avançados utilizando a análise univariada (ABULI et al. 2010). Vários estudos anteriores também verificaram a falta de correlação estatística entre SNPs e a sobrevida do CCR, exceto um realizado em chineses da etnia Han, no qual verificou-se a redução do risco de morrer e de recidiva tumoral entre os homozigotos comuns para os rs4779584 (SNP 26) e rs10795668 (SNP 59) (XING et al. 2011). As divergências entre os estudos chinês, norte-americano e europeu parecem ser mais de ordem metodológica do que por diferenças entre as frequências alélicas (HOSKINS et al. 2012). No presente estudo, também houve associação entre os homozigotos comuns dos SNPs 26 (genótipo CC) e 59 (genótipo GG) e o estadiamento precoce (0 e I, respectivamente). ABULI et al. (2010) também verificaram tal associação entre o SNP 59 e tumores precoces.

Neste estudo verificou-se uma chance maior de óbito por câncer entre os homozigotos raros dos SNPs 09 (genótipo TT) e 45 (genótipo CC). As variantes rs10411210 (SNP 09) e rs961253 (SNP 45) situam-se próximos dos genes *RHPN2* e *BMP2*, respectivamente, e também são membros da superfamília TGF- β (TENESA e DUNLOP 2009). No estudo escocês de TENESA et al. (2010), por outro lado, não houve correlação com nenhum dos dez SNPs e a mortalidade câncer-específica, mas com certa tendência quanto ao SNP 45, após ajustar por estadiamento, idade e sexo ($p = 0,07$). Como não houve associação estatisticamente significativa entre o risco de

CCR e os *SNPs* 45 e 59 no presente estudo, não é possível inferir que houve correlação direta com a mortalidade por câncer.

Além de influenciarem o risco de desenvolver o CCR, é provável que as variantes também influenciem a história natural da doença, tornando-se potenciais fatores prognósticos, como progressão e metástase. Entretanto, todos os estudos prognósticos foram retrospectivos (assim como este estudo) e se restringiram às mesmas variantes de predisposição, limitando, assim, seu impacto na prática clínica.

Em relação à idade, os homozigotos raros para o *SNP* 26 (genótipo TT) correlacionaram-se com indivíduos com 50 anos ou mais, enquanto que os heterozigotos para o *SNP* 82 (genótipo AC), com os mais jovens (menores que 50 anos). Entre os GWAS, o rs3802842 (*SNP* 82) também mostrou associação com casos mais jovens (abaixo dos 66 anos), enquanto que os *SNPs* 71 e 45 associaram-se com indivíduos mais velhos (HE et al. 2011). ABULI et al. (2010) também encontraram associação entre o *SNP* 82 e idade precoce ao diagnóstico (abaixo dos 50 anos). O rs16892766 (*SNP* 83) teve maior efeito nos casos abaixo dos 60 anos (TOMLINSON et al. 2008). Um estudo replicativo em suecos encontrou associações entre o *SNP* 71 e pacientes acima dos 60 anos, e entre homozigotos para os *SNPs* 09 e 59 e pacientes mais jovens (VON HOLST et al. 2010).

Em relação ao sexo, houve associação entre os homozigotos comuns para o *SNP* 45 (AA) e o sexo masculino, porém, sem relevância clínica, já que este *SNP* não obteve associação com o risco de CCR. Entre os GWAS e os estudos replicativos, um estudo revelou associação entre os *SNPs* 09 e

21 com o sexo masculino (HE et al. 2011), enquanto que outro encontrou o alelo de suscetibilidade do *SNP* 21 mais frequentemente nas mulheres (HOULSTON et al. 2008).

Quanto à localização do tumor, ABULI et al. (2010) encontraram associação entre o *SNP* 26 e tumores retais, e entre os *SNPs* 16 e 09 e tumores no cólon. Já, TOMLINSON et al. (2008) verificaram que o *SNP* 59 conferiu maior risco de tumores no reto do que no cólon. TENESA et al. (2008), por sua vez, encontraram esta correlação entre os *SNPs* 06 e 82. Interessantemente, nos indivíduos japoneses do estudo (7.579 no total), não houve aumento de risco para câncer de cólon em comparação com os europeus, evidenciando assim o primeiro alelo de suscetibilidade específico de uma população. Neste estudo, como houve apenas 60 indivíduos asiáticos, não foi possível revelar tal associação. O câncer retal foi mais comum entre os homocigotos para o *SNP* 45 (genótipo CC).

Em relação ao grau de diferenciação do tumor, apenas um estudo encontrou efeito maior dos *SNPs* 71 e 06 em tumores bem diferenciados ABULI et al. (2010). Neste estudo, porém, não foi encontrada correlação com tumores bem diferenciados, mas sim, entre os moderadamente e pouco diferenciados e os *SNPs* 83 (genótipo CC) e 06 (genótipo CT). Quanto ao tipo histológico, não foi encontrado nenhum estudo com essa variável na literatura. No presente estudo, foi encontrado com maior frequência o tipo tubular entre os heterocigotos para o *SNP* 09 (genótipo CT) com tendência à significância estatística ($p = 0,053$).

Em relação à história familiar, neste estudo houve correlação entre os homozigotos comuns (genótipo TT) do *SNP* 06 e um menor risco para história familiar até parentes de 3º grau. No GWAS original, no entanto, foram selecionados 940 indivíduos com CCR familiar nos quais foi identificado o *SNP* 06 (BRODERICK et al. 2007). Entre os outros GWAS e metanálises, nenhum encontrou correlação com a história familiar, e entre os estudos replicativos, somente três encontraram correlação entre os *SNPs* 06, 71 e 59, os dois primeiros *SNPs* com maior chance entre parente de 1º grau, e o último, entre homozigotos raros e câncer esporádico (TUUPANEN et al. 2008; ABULI et al. 2010; VON HOLST et al. 2010). Um estudo finlandês verificou a diferença entre os CCR esporádico e familiar e a correlação com o número de alelos de risco em 10 *SNPs* (os mesmos deste estudo). Observou-se que a média de alelos de risco nos casos familiares foi 11, enquanto que nos casos esporádicos foi 10, e para cada alelo, a chance de ter história familiar aumentou 16%. Os indivíduos com 15 alelos de risco tiveram mais que o dobro da chance de ter um parente de 1º grau afetado do que os indivíduos com 10 alelos (NIITTYMÄKI et al. 2010).

Neste estudo também se utilizaram os critérios revisados de Bethesda (UMAR et al. 2004) para avaliar a associação entre os alelos de risco e os CCR esporádico e familiar. Segundo um recente estudo populacional em mais de 2.000 casos com CCR, a sensibilidade e especificidade dos critérios de Bethesda para o diagnóstico da síndrome de Lynch é de 86% e 78%, respectivamente (PEREZ-CARBONELL et al. 2012). Verificou-se correlação entre os indivíduos heterozigotos para o *SNP* 82 (genótipo AC) com critérios

positivos, enquanto que os homozigotos raros (genótipo CC) para os *SNP* 16 e 26 correlacionaram-se com o tipo esporádico. Nos GWAS originais, não foi observada correlação desses *SNPs* com as variáveis pesquisadas, porém nenhum utilizou os critérios de Bethesda como variável (HOULSTON et al. 2008; TENESA et al. 2008; JAEGER et al. 2008). Apenas dois estudos utilizaram a pesquisa de instabilidade de microssatélites para a correlação, e um deles encontrou associação entre os tumores estáveis e o *SNP* 16, de certa forma, favorecendo ao câncer esporádico, já que a minoria dos tumores esporádicos apresentam instabilidade de microssatélites (HOULSTON et al. 2008; TOMLINSON et al. 2008).

Em relação ao risco de adenoma avançado em relação ao de CCR, os homozigotos raros para o *SNP* 26 (genótipo TT) apresentaram maior risco de adenomas. Apenas um estudo verificou tal associação, porém com os *SNPs* 21 e 71, ambos em homozigose (ABULI et al. 2010). JAEGER et al. (2008) identificaram o *SNP* 26 (rs 4779584) como fator de risco para o CCR, embora não tenham observado diferença na formação de pólipos serrilhados entre os indivíduos com o alelo de risco. Esta variante situa-se próxima ao gene *GREM1*, um antagonista dos genes BMP, membros da família TGF- β (TENESA e DUNLOP 2009). A perda dos sinais em BMP foi relatada na progressão de adenomas avançados para câncer precoce (HARDWICK et al. 2008).

5.5 CONSIDERAÇÕES FINAIS

Apesar dos avanços na compreensão da etiologia da suscetibilidade genética ao CCR, desafios consideráveis ainda permanecem, tais como:

- identificar as variantes causais responsáveis pelas associações;
- elucidar as consequências funcionais das variantes causais;
- determinar a contribuição dessas variantes para o risco genético em outras populações.

Outro desafio é identificar parte da “herdabilidade faltante” do CCR atribuída às variantes raras causais e com maior penetrância. A última geração de sequenciadores de alta *performance* tem reduzido os custos do ressequenciamento do genoma com uma grande abrangência e oferece o potencial para identificar tais variantes. Obviamente, devem também ser incluídos os fatores ambientais nas estimativas de risco, já que contribuem com a maior parte da etiologia do CCR e que, pelo menos em teoria, mas provavelmente na prática, algumas variantes possam ter maiores ou menores efeitos, dependendo dos fatores do ambiente (dieta, estilo de vida, consumo de álcool, etc.). Nesse sentido, a abordagem com vários estudos colaborativos e consórcios internacionais que englobem diferentes populações, ajudará a elucidar as interações entre as variantes genéticas e os fatores ambientais.

Por fim, considerando as potenciais implicações de tais pesquisas em medidas de saúde pública, no sentido de prevenir o CCR e reduzir sua incidência e mortalidade, seja pelo diagnóstico e tratamento precoces de

pólipos adenomatosos por exames de rastreamento, como a colonoscopia, seja pela quimioprofilaxia com o uso da aspirina, por exemplo, a estratificação do risco populacional poderia ser uma medida custo-efetiva, na medida em que os recursos sejam empregados com equidade sócio-econômica.

6 CONCLUSÃO

O presente estudo de associação caso-controle procurou identificar na população brasileira dez variantes comuns de suscetibilidade ao CCR que haviam sido previamente identificadas em populações de ancestralidade europeia. Foi obtida associação estatisticamente significativa com o risco de CCR em metade dos *SNPs* pesquisados, a saber:

rs4939827 ou *SNP* 06;

rs10411210 ou *SNP* 09;

rs4444235 ou *SNP* 16;

rs3802842 ou *SNP* 82 e

rs16892766 ou *SNP* 83

Os alelos de risco de CCR mostraram frequências próximas às dos europeus, assim como os efeitos de cada alelo/genótipo também foram modestos. O *SNP* 06 foi a variante que obteve o maior efeito sobre o risco de CCR e obteve a associação mais significativa, conferindo, portanto o maior risco atribuível populacional.

No que se refere às correlações genótipo-fenótipo, foram encontradas associações entre todas as variáveis analisadas, com destaque para o *SNP* 06 apontando para um melhor prognóstico do CCR e o *SNP* 82 com os critérios revisados de Bethesda. Entretanto, deve-se salientar que as análises foram univariadas e retrospectivas.

7 REFERÊNCIAS BIBLIOGRÁFICAS

Aaltonen L, Johns L, Jarvinen H, Mecklin JP, Houlston R. Explaining the familial colorectal cancer risk associated with mismatch repair (MMR)-deficient and MMR-stable tumors. **Clin Cancer Res** 2007; 13:356-61.

Aaltonen LA, Salovaara R, Kristo P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. **N Engl J Med** 1998; 338:1481-7.

Abuli A, Bessa X, Gonzalez JR, et al. Susceptibility genetic variants associated with colorectal cancer risk correlate with cancer phenotype. **Gastroenterology** 2010; 139:788-96.

Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. **Nat Genet** 2006; 38:652-8.

Antoniou AC, Easton DF. Polygenic inheritance of breast cancer: Implications for design of association studies. **Genet Epidemiol** 2003; 25:190-202.

Atkin WS, Edwards R, Kralj-Hans I, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. **Lancet** 2010; 375:1624-33.

Balding DJ. A tutorial on statistical methods for population association studies. **Nat Rev Genet** 2006; 7:781-91.

Barnholtz-Sloan JS, Raska P, Rebbeck TR, Millikan RC. Replication of GWAS "Hits" by Race for Breast and Prostate Cancers in European Americans and African Americans. **Front Genet** 2011; 2:37.

Berndt SI, Potter JD, Hazra A, et al. Pooled analysis of genetic variation at chromosome 8q24 and colorectal neoplasia risk. **Hum Mol Genet** 2008; 17:2665-72.

Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. **Nat Genet** 2008; 40:695-701.

Bost B, de Vienne D, Hospital F, Moreau L, Dillmann C. Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. **Genetics** 2001; 157:1773-87.

Broderick P, Carvajal-Carmona L, Pittman AM, et al. A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. **Nat Genet** 2007; 39:1315-7.

Cannon-Albright LA, Skolnick MH, Bishop DT, Lee RG, Burt RW. Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. **N Engl J Med** 1988; 319:533-7.

Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. **Cancer Epidemiol Biomarkers Prev** 2009; 18:1688-94.

Cunningham D, Atkin W, Lenz HJ, et al. Colorectal cancer. **Lancet** 2010; 375:1030-47.

Cunningham JM, Christensen ER, Tester DJ, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. **Cancer Res** 1998; 58:3455-60.

Curtin K, Lin WY, George R, et al. Meta association of colorectal cancer confirms risk alleles at 8q24 and 18q21. **Cancer Epidemiol Biomarkers Prev** 2009; 18:616-21.

de Jong AE, Hendriks YM, Kleibeuker JH, et al. Decrease in mortality in Lynch syndrome families because of surveillance. **Gastroenterology** 2006; 130:665-71.

Dickson SP, Wang K, Krantz I, et al. Rare variants create synthetic genome-wide associations. **PLoS Biol** 2010; 8:e1000294.

Dunlop MG, Dobbins SE, Farrington SM, et al. Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. **Nat Genet** 2012a; 44:770-6.

Dunlop MG, Tenesa A, Farrington SM, et al. Cumulative impact of common genetic variants and other risk factors on colorectal cancer risk in 42,103 individuals. **Gut** 2012b May 22 [Epub ahead of print].

Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. **Nature** 2007; 447:1087-93.

Edwards BK, Ward E, Kohler BA, et al. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. **Cancer** 2010; 116:544-73.

Ehret GB, Lamparter D, Hoggart CJ, et al. A Multi-SNP locus-association method reveals a substantial fraction of the missing heritability. **Am J Hum Genet** 2012; 91:863-871.

[EGAPP] Evaluation of Genomic Applications in Practice and Prevention Working G. Recommendations from the EGAPP Working Group: genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. **Genet Med** 2009; 11:35-41.

Evans DM, Purcell S. Power calculations. In: Weiner MP, Gabriel SB, Stephens JC, editors. **Genetic variation a laboratory manual**. New York: Cold Spring Harbor Laboratory Press; 2007. p.19-30.

Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. **Behav Res Methods** 2009; 41:1149-60.

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. **Int J Cancer** 2010; 127:2893-917.

Fletcher O, Houlston RS. Architecture of inherited susceptibility to common cancer. **Nat Rev Cancer** 2010; 10:353-61.

Giolo SR, Soler JM, Greenway SC, et al. Brazilian urban population genetic structure reveals a high degree of admixture. **Eur J Hum Genet** 2012; 20:111-6.

Haiman CA, Le Marchand L, Yamamoto J, et al. A common genetic risk factor for colorectal and prostate cancer. **Nat Genet** 2007; 39:954-6.

Hardwick JC, Kodach LL, Offerhaus GJ, van den Brink GR. Bone morphogenetic protein signalling in colorectal cancer. **Nat Rev Cancer** 2008; 8:806-12.

He J, Wilkens LR, Stram DO, et al. Generalizability and epidemiologic characterization of eleven colorectal cancer GWAS hits in multiple populations. **Cancer Epidemiol Biomarkers Prev** 2011; 20:70-81.

Hoskins JM, Ong PS, Keku TO, et al. Association of eleven common, low-penetrance colorectal cancer susceptibility genetic variants at six risk loci with clinical outcome. **PloS One** 2012; 7:e41954.

Houlston RS, Webb E, Broderick P, et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. **Nat Genet** 2008; 40:1426-35.

Houlston RS, Cheadle J, Dobbins SE, et al. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. **Nat Genet** 2010; 42:973-7.

International HapMap Consortium. A haplotype map of the human genome. **Nature** 2005; 437:1299-320.

Jaeger EE, Woodford-Richens KL, Lockett M, et al. An ancestral Ashkenazi haplotype at the HMPS/CRAC1 locus on 15q13-q14 is associated with hereditary mixed polyposis syndrome. **Am J Hum Genet** 2003; 72:1261-7.

Jaeger E, Webb E, Howarth K, et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. **Nat Genet** 2008; 40:26-8.

Kupfer SS, Torres JB, Hooker S, et al. Novel single nucleotide polymorphism associations with colorectal cancer on chromosome 8q24 in African and European Americans. **Carcinogenesis** 2009; 30:1353-7.

Laken SJ, Petersen GM, Gruber SB, et al. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. **Nat Genet** 1997; 17:79-83.

Lee LG, Connell CR, Bloch W. Allelic discrimination by nick-translation PCR with fluorogenic probes. **Nucleic Acids Res** 1993; 21:3761-6.

Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. **Nature** 1998; 396:643-9.

Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. **Gastroenterology** 2008; 134:1570-95.

Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. **N Engl J Med** 2000; 343:78-85.

Matsuo K, Suzuki T, Ito H, et al. Association between an 8q24 locus and the risk of colorectal cancer in Japanese. **BMC Cancer** 2009; 9:379.

Middeldorp A, Jagmohan-Changur S, van Eijk R, et al. Enrichment of low penetrance susceptibility loci in a Dutch familial colorectal cancer cohort. **Cancer Epidemiol Biomarkers Prev** 2009; 18:3062-7.

Ministério da Saúde. Instituto Nacional de Câncer. **Estimativa/2012: incidência de câncer no Brasil**. Rio de Janeiro: INCA; 2011.

National Human Genome Research Institute. **Division of genomic medicine a catalog of published genome-wide association studies**. Available from: <URL:<http://www.genome.gov/26525384>> [2012 Nov 20].

Need AC, Goldstein DB. Next generation disparities in human genomics: concerns and remedies. **Trends Genet** 2009; 25:489-94.

Niittymaki I, Kaasinen E, Tuupanen S, et al. Low-penetrance susceptibility variants in familial colorectal cancer. **Cancer Epidemiol Biomarkers Prev** 2010; 19:1478-83.

Noffsinger AE. Serrated polyps and colorectal cancer: new pathway to malignancy. **Annu Rev Pathol** 2009; 4:343-64.

Pena SD, Di Pietro G, Fuchshuber-Moraes M, et al. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. **PloS One** 2011; 6:e17063.

Perez-Carbonell L, Ruiz-Ponte C, Guarinos C, et al. Comparison between universal molecular screening for Lynch syndrome and revised Bethesda guidelines in a large population-based cohort of patients with colorectal cancer. **Gut** 2012; 61:865-72.

Peters U, Hutter CM, Hsu L, et al. Meta-analysis of new genome-wide association studies of colorectal cancer risk. **Hum Genet** 2012; 131:217-34.

Pomerantz MM, Ahmadiyeh N, Jia L, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. **Nat Genet** 2009; 41:882-4.

Poynter JN, Figueiredo JC, Conti DV, et al. Variants on 9p24 and 8q24 are associated with risk of colorectal cancer: results from the Colon Cancer Family Registry. **Cancer Res** 2007; 67:11128-32.

Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. **Nature** 2001; 409:928-33.

Schoen RE, Pinsky PF, Weissfeld JL, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. **N Engl J Med** 2012; 366:2345-57.

Shriner D, Adeyemo A, Ramos E, Chen G, Rotimi CN. Mapping of disease-associated variants in admixed populations. **Genome Biol** 2011; 12:223.

St John DJ, McDermott FT, Hopper JL, Debney EA, Johnson WR, Hughes ES. Cancer risk in relatives of patients with common colorectal cancer. **Ann Intern Med** 1993; 118:785-90.

Stadler ZK, Thom P, Robson ME, et al. Genome-wide association studies of Cancer. **J Clin Oncol** 2010; 28:4255-4267.

Strachan T, Read A. **Human molecular genetics**. 4^a ed. New York: Garland Science; 2011. Human genetic variability and its consequences; p.405-40.

Tenesa A, Farrington SM, Prendergast JG, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. **Nat Genet** 2008; 40:631-7.

Tenesa A, Dunlop MG. New insights into the aetiology of colorectal cancer from genome-wide association studies. **Nat Rev Genet** 2009; 10:353-8.

Tenesa A, Theodoratou E, Din FV, et al. Ten common genetic variants associated with colorectal cancer risk are not associated with survival after diagnosis. **Clin Cancer Res** 2010; 16:3754-9.

Thompson CL, Plummer SJ, Acheson LS, Tucker TC, Casey G, Li L. Association of common genetic variants in SMAD7 and risk of colon cancer. **Carcinogenesis** 2009; 30:982-6.

Tomlinson I, Webb E, Carvajal-Carmona L, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. **Nat Genet** 2007; 39:984-8.

Tomlinson IP, Webb E, Carvajal-Carmona L, et al. A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. **Nat Genet** 2008; 40:623-30.

Tomlinson IP, Dunlop M, Campbell H, et al. COGENT (COlorectal cancer GENEtics): an international consortium to study the role of polymorphic variation on the risk of colorectal cancer. **Br J Cancer** 2010; 102:447-54.

Tomlinson IP, Carvajal-Carmona LG, Dobbins SE, et al. Multiple common susceptibility variants near BMP pathway loci GREM1, BMP4, and BMP2 explain part of the missing heritability of colorectal cancer. **PLoS Genet** 2011; 7:e1002105.

Tuupanen S, Niittymaki I, Nousiainen K, et al. Allelic imbalance at rs6983267 suggests selection of the risk allele in somatic colorectal tumor evolution. **Cancer Res** 2008; 68:14-7.

Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. **J Natl Cancer Inst** 2004; 96:261-8.

Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. **Am J Hum Genet** 2012; 90:7-24.

Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. **N Engl J Med** 1988; 319:525-32.

von Holst S, Picelli S, Edler D, et al. Association studies on 11 published colorectal cancer risk loci. **Br J Cancer** 2010; 103:575-80.

Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. **Nature** 2007; 447:661-78.

Whitelaw SC, Murday VA, Tomlinson IP, et al. Clinical and molecular features of the hereditary mixed polyposis syndrome. **Gastroenterology** 1997; 112:327-34.

Xing J, Myers RE, He X, et al. GWAS-identified colorectal cancer susceptibility locus associates with disease prognosis. **Eur J Cancer** 2011; 47:1699-707.

Xiong F, Wu C, Bi X, et al. Risk of genome-wide association study-identified genetic variants for colorectal cancer in a Chinese population. **Cancer Epidemiol Biomarkers Prev** 2010; 19:1855-61.

Zanke BW, Greenwood CM, Rangrej J, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. **Nat Genet** 2007; 39:989-94.

Anexo 1 - Termo de Consentimento Livre e Esclarecido para Coleta de Sangue (Casos)

Termo de Consentimento Livre e Esclarecido para Indivíduos com Câncer Colorretal (Conforme os termos das Resoluções 196/96 e CNS 251/97 do Ministério da Saúde)

NOME DO PROJETO: **“Estudo Genético do Câncer Colorretal na Hispânia e nas Américas”**

I. DADOS DE IDENTIFICAÇÃO DO PACIENTE OU RESPONSÁVEL LEGAL

NOME DO PARTICIPANTE: _____

Sexo: _____ Data de nascimento: ____/____/____

Documento de identidade n°: _____

Endereço: _____ Complemento: _____

Cidade: _____ Estado: _____

CEP: _____ TEL: _____

RESPONSÁVEL (se menor ou incapaz) : _____

Sexo: _____ Data de nascimento: ____/____/____

Documento de identidade n.: _____

Endereço: _____

Número: _____ Complemento: _____ Cidade: _____ Estado: _____

II. OBJETIVOS DO ESTUDO

Você foi convidado a participar deste estudo, pois você teve câncer de intestino grosso. O objetivo deste estudo é pesquisar alterações genéticas que possam estar relacionadas ao risco de desenvolver câncer no intestino grosso.

III. PROCEDIMENTOS

Se você possuir amostra de sangue armazenada no Hospital, você será informado por telefone, e para participar, basta que assine esta autorização. A partir dos dados que estão no seu prontuário será preenchido um formulário com características como sexo, idade de diagnóstico, tipo de tumor, e tratamento realizado.

Se você não possui amostra de sangue armazenada, você receberá um telefonema para agendar a coleta de sangue. Esta coleta deverá coincidir com uma de suas visitas ao Hospital. No dia da coleta você será encaminhado ao laboratório para a retirada da amostra

de sangue. No laboratório serão retirados 8ml de sangue (aproximadamente uma colher de sopa) através de uma agulha colocada na veia do seu braço. O sangue será colocado em dois tubos, que serão identificados com um código e enviados ao laboratório. Nesta pesquisa serão realizados testes que irão buscar por alterações genéticas que possam aumentar o risco de câncer de intestino grosso. As características clínicas e os resultados dos testes na amostra de sangue dos participantes com câncer serão comparadas com as características de participantes sem câncer.

Os resultados deste estudo serão utilizados para apresentação de artigos e trabalhos científicos. Na divulgação dos resultados não serão divulgados os nomes ou qualquer forma que possa identificar os participantes da pesquisa.

Também solicitamos sua autorização para guardar uma parte da amostra de sangue (um tubo) que não será utilizada nesta pesquisa. Esta amostra será guardada no banco de tumores do Hospital A.C. Camargo, e poderá ser utilizada em pesquisas futuras. Todas as pesquisas são aprovadas pelo Comitê de Ética em Pesquisa antes do seu início, e se necessário solicitaremos sua autorização para uso. Você pode decidir participar da pesquisa, mas pode não permitir que a amostra seja guardada após o final desta pesquisa.

IV - BENEFÍCIOS

A sua participação neste estudo pode não trazer benefícios diretamente para você ou para os membros de sua família, mas pode ajudar os médicos a compreender como o câncer se desenvolve. Não será possível avaliar o risco de câncer da sua família com estes resultados, pois as alterações que podemos encontrar são desconhecidas, e depois de encontradas, serão necessários outros estudos para verificar o papel destas alterações.

V - RISCOS

Os riscos associados com sua participação nesse estudo são dor ou queimação no local da retirada do sangue.

VI – ASPECTOS ÉTICOS

Esta pesquisa foi aprovada por um Comitê de Ética. A confidencialidade das suas informações será mantida. Apenas as pessoas envolvidas diretamente nesse estudo ou o Comitê de Ética em Pesquisa poderão ter acesso às informações, se necessário.

A sua participação neste estudo é voluntária, tendo o direito de retirar-se do estudo a qualquer momento. Sua recusa ou desistência não irá prejudicar o seu tratamento.

Qualquer dúvida sobre o estudo, você poderá entrar em contato com o pesquisador principal (Dr.Samuel Aguiar Jr.) e/ou pessoa por mim delegada no telefone 2189-5000 ramal 1080. Se o pesquisador principal não fornecer as informações/esclarecimentos suficientes, entre em contato com o Coordenador do Comitê de Ética do Hospital do Câncer – SP, pelo telefone 21895020.

Declaro que fui esclarecido: sobre os procedimentos, riscos e benefícios sobre este estudo; que tenho liberdade em retirar o meu consentimento a qualquer momento, sem que isto traga prejuízo a continuidade do tratamento; que não haverá remuneração financeira para este estudo; sobre a segurança de que minha identidade será preservada, mantendo-se todas as informações em caráter confidencial.

() concordo em participar do estudo

() concordo com a guarda da amostra de sangue

São Paulo, _____ de _____ de _____

Assinatura do Participante ou Responsável

Assinatura do Pesquisador ou Representante

Anexo 2 - Termo de Consentimento Livre e Esclarecido para Coleta de Sangue (Controles)

Termo de Consentimento Livre e Esclarecido para Indivíduos sem Câncer Colorretal

(Conforme os termos das Resoluções 196/96 e CNS 251/97 do Ministério da Saúde)

NOME DO PROJETO: “Estudo Genético de Câncer Colorretal na Hispânia e nas Américas”

I. DADOS DE IDENTIFICAÇÃO DO PACIENTE OU RESPONSÁVEL LEGAL

NOME DO PARTICIPANTE: _____

Sexo: _____ Data de nascimento: ___/___/___

Documento de identidade n.: _____

Endereço: _____ Número: _____ Complemento: _____

Cidade: _____ Estado: _____

CEP: _____ TEL: _____

RESPONSÁVEL: _____

Sexo: _____ Data de nascimento: ___/___/___

Documento de identidade n.: _____

Endereço: _____

Número: _____ Complemento: _____

Cidade: _____ Estado: _____

CEP: _____ TEL: _____

II. OBJETIVOS DO ESTUDO

O objetivo deste estudo é pesquisar por alterações genéticas que possam estar relacionadas a um aumento no risco de câncer do intestino grosso.

III. PROCEDIMENTOS

Após a assinatura deste termo, será preenchido um formulário com características como sexo, idade e história familiar. Depois você será encaminhado ao laboratório para a retirada da amostra de sangue. No laboratório serão retirados 8ml de sangue (aproximadamente uma colher de sopa) através de uma agulha colocada no seu braço. O sangue será colocado em dois tubos, identificado e enviado ao laboratório para análise. Um dos tubos de sangue ficará no Brasil e o outro tubo será enviado para a Inglaterra para a realização da pesquisa. Nesta pesquisa será realizado um teste que irá buscar alterações genéticas que

possam aumentar o risco de câncer de intestino grosso. As características clínicas e os resultados dos testes na amostra de sangue dos participantes com câncer será comparado com as características de participantes sem câncer.

Os resultados deste estudo serão utilizados para apresentação de artigos e trabalhos científicos. Na divulgação dos resultados não serão divulgados os nomes ou qualquer forma que possa identificar os participantes da pesquisa.

Também solicitamos sua autorização para guardar parte da amostra de sangue (dois tubos) que não será utilizada nesta pesquisa. Esta amostra será guardada no banco de tumores do Hospital A.C. Camargo, e poderá ser utilizada em pesquisas futuras. Todas as pesquisas são aprovadas pelo Comitê de Ética em Pesquisa antes do seu início, e se necessário solicitaremos sua autorização para uso. Você pode decidir participar da pesquisa, mas não permitir que a amostra seja guardada após o final desta pesquisa.

IV - BENEFÍCIOS

A participação estudo pode não trazer benefícios para você ou para os membros de sua família; mas pode ajudar os médicos a compreender como o câncer se desenvolve.

V - RISCOS

Os riscos associados com sua participação nesse estudo são dor ou queimação no local da retirada do sangue.

VI – ASPECTOS ÉTICOS

Esta pesquisa foi aprovada por um Comitê de Ética. A confidencialidade das suas informações será mantida. Apenas as pessoas envolvidas diretamente nesse estudo ou o Comitê de Ética em Pesquisa poderão ter acesso as informações se necessário.

A sua participação neste estudo é voluntária, tendo o direito de retirar-se do estudo a qualquer momento. Sua recusa ou desistência não irá prejudicar o tratamento do seu familiar.

Qualquer dúvida sobre o estudo, você poderá entrar em contato com o pesquisador principal (Samuel Aguiar Jr.) e/ou pessoa por mim delegada no telefone 2189-5000 ramal 1080. Se o pesquisador principal não fornecer as informações/esclarecimentos suficientes, entre em contato com o Coordenador do Comitê de Ética do Hospital do Câncer – SP, pelo telefone 21895020.

Declaro que fui esclarecido: sobre os procedimentos, riscos e benefícios sobre este estudo; que tenho liberdade em retirar o meu consentimento a qualquer momento; que não haverá remuneração financeira para este estudo; sobre a segurança de que minha identidade será preservada, mantendo-se todas as informações em caráter confidencial.

() concordo participar do estudo

() concordo com a guarda da amostra de sangue

São Paulo, _____ de _____ de _____

Assinatura do Participante ou Responsável

Assinatura do Pesquisador ou Representante

Anexo 3 - Termo de Consentimento Livre e Esclarecido para Amostras do Banco de Macromoléculas (Casos)

Termo de Consentimento Livre e Esclarecido para Indivíduos com Câncer Colorretal (Conforme os termos das Resoluções 196/96 e CNS 251/97 do Ministério da Saúde)

NOME DO PROJETO: **“Estudo Genético de Câncer Colorretal na Hispânia e nas Américas”**

I. DADOS DE IDENTIFICAÇÃO DO PACIENTE OU RESPONSÁVEL LEGAL

NOME DO PARTICIPANTE: _____

Sexo: _____ Data de nascimento: ____/____/____

Documento de identidade (RG ou CPF): _____

Endereço: _____ Complemento: _____

Cidade: _____ Estado: _____

CEP: _____ TEL: _____

RESPONSÁVEL (se menor ou incapaz): _____

Sexo: _____ Data de nascimento: ____/____/____

Documento de identidade (RG ou CPF): _____

Endereço: _____

Complemento: _____

Cidade: _____ Estado: _____

CEP: _____ TEL: _____

II. OBJETIVOS DO ESTUDO

Você foi convidado(a) a participar deste estudo, pois você teve câncer de intestino grosso. O objetivo deste estudo é pesquisar alterações genéticas que possam estar relacionadas ao risco de desenvolver câncer no intestino grosso.

III. PROCEDIMENTOS

Se você possuir amostra de sangue armazenada no Hospital (seu caso), você será (foi) informado(a) por telefone, e para participar, basta que assine esta autorização. A partir dos dados que estão no seu prontuário será preenchido um formulário com características como sexo, idade de diagnóstico, tipo de tumor, e tratamento realizado.

Os resultados deste estudo serão utilizados para apresentação de artigos e trabalhos científicos. Na divulgação dos resultados não serão divulgados os nomes ou qualquer forma que possa identificar os participantes da pesquisa.

IV - BENEFÍCIOS

A sua participação neste estudo pode não trazer benefícios diretamente para você ou para os membros de sua família, mas pode ajudar os médicos a compreender como o câncer se desenvolve. Não será possível avaliar o risco de câncer na sua família com estes resultados, pois as alterações que podemos encontrar são desconhecidas, e depois de encontradas, serão necessários outros estudos para verificar o papel destas alterações.

V – ASPECTOS ÉTICOS

Esta pesquisa foi aprovada por um Comitê de Ética. A confidencialidade das suas informações será mantida. Apenas as pessoas envolvidas diretamente nesse estudo ou o Comitê de Ética em Pesquisa poderão ter acesso às informações, se necessário.

A sua participação neste estudo é voluntária, tendo o direito de retirar-se do estudo a qualquer momento. Sua recusa ou desistência não irá prejudicar o seu tratamento.

Qualquer dúvida sobre o estudo, você poderá entrar em contato com o pesquisador principal (Dr. Samuel Aguiar Jr) e/ou pessoa por mim delegada no telefone 2189-5000 ramal 1080. Se o pesquisador principal não fornecer as informações/esclarecimentos suficientes, entre em contato com o Coordenador do Comitê de Ética do Hospital do Câncer – SP, pelo telefone 21895020.

Declaro que fui esclarecido sobre: os procedimentos e benefícios deste estudo; a liberdade de retirar o meu consentimento a qualquer momento, sem que isto traga prejuízo à continuidade do meu tratamento; que não haverá remuneração financeira para este estudo; a segurança de que minha identidade será preservada, mantendo-se todas as informações em caráter confidencial.

() concordo em participar do estudo

São Paulo, _____ de _____ de _____

Assinatura do Participante ou Responsável

Assinatura do Pesquisador Responsável ou Representante

Anexo 4 - Ficha para Coleta de Dados Clínico-Patológicos e História Familiar

	Nome	
	RGH	
	Data de nascimento:	Data óbito:
	Sexo (1) Masculino (2) Feminino	Etnia:
	Data de inclusão no estudo	
	Classificação segundo história familiar (heredograma no verso)	S N Ign
	Esporádico	□□□□□□□□
	HNPCC (Amsterdam I).....	□□□□□□□□
	HNPCC (Amsterdam II).....	□□□□□□□□
	Bethesda.....	□□□□□□□□
	Agrupamento Familiar.....	□□□□□□□□
	FAP.....	□□□□□□□□
	AFAP.....	□□□□□□□□
	Outras poliposes	□□□□□□□□
	Tumores	S N Idade
		Dx□
	Colorretal (exceto adenomas).....	□□□□□_____
	Adenomas colorretais.....	□□□□□_____
	Endométrio	□□□□□_____
	Estômago.....	□□□□□_____
	Sistema hepatobiliar.....	□□□□□_____□
	Intestino delgado.....	□□□□□_____□
	Pelve renal e/ou ureter	□□□□□_____□
	Ovário	□□□□□_____
	Pâncreas	□□□□□_____
	SNC.....	□□□□□_____
	Mama	□□□□□_____□
	Outros	□□□□□_____□
	Tumor primário colorretal	
	(0) não tem CCR (1) cólon prox (D+T) (2) cólon distal (E+S) (3) reto	
	(4) adenomas (9) ignorado	
	Estadiamento clínico:	
	Estádio T (quando possível usar estágio T patológico)	tamanho
	(0) T0 (1) T1 (2) T2 (3) T3	
	(4) T4 (9) Tx	
	Estádio N (quando possível usar estágio N patológico)	número
	(0) N0 (1) N1 (2) N2 (9) Nx	
	Estádio M (quando possível usar estágio M patológico)	
	(0) M0 (1) M1 retroperitônio (2) M1 fígado (3) M1 pulmão	
	(4) M1 outro (_____) (9) Mx	
	Grau de diferenciação	
	(1) Bem Diferenciado (2) moderadamente diferenciado (3) pouco diferenciado (4) indiferenciado (8) outro (9) ign	
nºAP	Tipo histológico	
	(1) Adenocarcinoma tubular (2) Adenoca com dif. Mucinoso (até 50%)	
	(3) Adenoca mucinoso (>50%) (4) Diferenciação neuroendócrina	
	(5) CEC (6) Cribriforme (<50%)	
	(7) Adenoescamoso (8) Indiferenciado	
	(9) Pequenas células (10) Anel de sinete	
	(11) Medular (12) outros (99) ign	
	HISTÓRIA FAMILIAL DE CCR () 1º GRAU () 2º GRAU () 3º GRAU	